



Herausgegeben vom Herausgeberverein
Zeitschrift Psychotherapie-Wissenschaft

Psychotherapie Wissenschaft

 WWW.PSYCHOTHERAPIE-WISSENSCHAFT.INFO

KI in der Psychotherapie

Herausgegeben von Mario Schlegel & Peter Schulthess

16. Jahrgang

Heft 2 / 2026

eISSN 1664-9591

Paolo Raile

KI in der Psychotherapie: Ersatz, Werkzeug oder Projektionsfläche?

Technische Grundlagen, ethische Implikationen, klinische Perspektiven

Jan M. Schürmann, Jana Sedlakova & Manuel Trachsel

Ethische Aspekte zur KI-Anwendung in der Psychotherapie

Jean Schulthess Watt

Adoleszenz, Identität und Social Media

Mario Schlegel

KI als Sennenpuppe, vertrautes Gegenüber und Werkzeug mit geistiger Rückwirkung

Essay

Frank Schulz-Kindermann

Psychotherapie bei fortgeschritten erkrankten Krebspatient:innen

Marie-Christine Hartlieb

Zur Modulation autonomer Regulationsprozesse bei ME/CFS innerhalb der Psychotherapie

Tamara M. Trebes

Zur epistemischen Eigenständigkeit der Psychotherapiewissenschaft

Plädoyer für eine Ethik der reflexiven Erfahrung



Psychosozial-Verlag



Psychotherapie Wissenschaft

🔗 WWW.PSYCHOTHERAPIE-WISSENSCHAFT.INFO

KI in der Psychotherapie

Herausgegeben von Mario Schlegel & Peter Schulthess

16. Jahrgang

Heft 2 / 2026

eISSN 1664-9591



Psychosozial-Verlag

Impressum

Psychotherapie-Wissenschaft

eISSN 1664-9591

16. Jahrgang Heft 2/2026

<https://doi.org/10.30820/1664-9583-2026-2>

info@psychotherapie-wissenschaft.info

www.psychotherapie-wissenschaft.info

Herausgeber

Herausgeberverein der Zeitschrift

Psychotherapie-Wissenschaft

Peter Schulthess

Zanggerweg 33

CH-8006 Zürich

Mobile + 41 76 559 19 20

E-Mail: peter.schulthess@psychotherapie-wissenschaft.info

Redaktion

Mara Foppoli, Lugano

Lea-Sophie Richter, Zürich

Mario Schlegel, Zürich

Peter Schulthess, Zürich

Nataliya Tereshchenko, Zürich/Wien

Hinweise für AutorInnen befinden
sich auf der Homepage der Zeitschrift:
www.psychotherapie-wissenschaft.info

Verlag

Psychosozial-Verlag GmbH & Co. KG

Walltorstr. 10

D-35390 Gießen

+49 6421 96 99 78 26

info@psychosozial-verlag.de

www.psychosozial-verlag.de

Anzeigen

Anfragen zu Anzeigen bitte an den Verlag:

anzeigen@psychosozial-verlag.de

Es gelten die Preise der auf www.psychosozial-verlag.de
einsehbaren Mediadaten.

Gönnermitglieder des Herausgebervereins

wenden sich bitte direkt an die Redaktion:

redaktion@psychotherapie-wissenschaft.info

Datenbanken

Die Zeitschrift Psychotherapie-Wissenschaft wird regelmäßig
in der Abstract-Datenbank PsycInfo der American Psycho-
logical Association (APA), im Directory of Open Journals
(DOAJ) und in den Publikationsdatenbanken PSYINDEX und
PsychArchives des Leibniz-Institut für Psychologie/Leibniz
Institute for Psychology (ZPID) erfasst.



Die Beiträge dieser Zeitschrift sind unter der Creative Commons
Attribution-NonCommercial-NoDerivs 3.0 DE Lizenz lizenziert.
Diese Lizenz erlaubt die private Nutzung und unveränderte
Weitergabe, verbietet jedoch die Bearbeitung und kommerzielle
Nutzung. Weitere Informationen finden Sie unter:
creativecommons.org/licenses/by-nc-nd/3.0/de

Inhalt

Editorial

Mario Schlegel & Peter Schulthess

Titelthema:

KI in der Psychotherapie

KI in der Psychotherapie:

Ersatz, Werkzeug oder Projektionsfläche?

Technische Grundlagen, ethische Implikationen,
klinische Perspektiven

Paolo Raile

Ethische Aspekte zur KI-Anwendung in der Psychotherapie

*Jan M. Schürmann, Jana Sedlakova &
Manuel Trachsel*

Adoleszenz, Identität und Social Media

Jean Schulthess Watt

Essay

KI als Sennenpuppe, vertrautes Gegenüber
und Werkzeug mit geistiger Rückwirkung

Essay

Mario Schlegel

5 Originalarbeiten

Psychotherapie bei fortgeschritten
erkrankten Krebspatient:innen

Frank Schulz-Kindermann

43

7 Zur Modulation autonomer Regulationsprozesse
bei ME/CFS innerhalb der Psychotherapie

Marie-Christine Hartlieb

49

Zur epistemischen Eigenständigkeit
der Psychotherapiewissenschaft

Plädoyer für eine Ethik der reflexiven Erfahrung

17 *Tamara M. Trebes*

55

Buchbesprechungen

29 Geuter, U. (2026). *Wege zum Körper.*
Schriften zur Körperpsychotherapie

Thomas Fellmann

63

Stöckli, P. (2026). *Autismus, Trauma und Bindung.*
Neue Wege zu Regulation und Verbindung

35 *Peter Schulthess*

64

Hell, D. (2026). *Fragile Psyche.*

Von Umbrüchen und Aufbrüchen

Peter Schulthess

65

Kongresskalender

66

Editorial

Heute dringt die «Künstliche Intelligenz», wie wir erstaunt feststellen müssen, in sämtliche Lebensbereiche ein. Es ist nicht wie damals, im 18. und 19. Jahrhundert, als die Wärmekraftmaschinen erfunden und die Elektrizität entdeckt wurde. Durch die schier ins Unendliche gesteigerte Muskelkraft durch die Maschinen, den Segen der Elektrizität und die elektronische Informationsverarbeitung hat die moderne Technik unsere Kultur bereichert, obwohl sie, wie wir heute wissen, die Umwelt erheblich belastet. Nicht zuletzt ist es auch die geistige Freiheit, die vom Fortschritt profitiert hat. Ganz anders verhält es sich bei der Potenzierung geistiger Fähigkeiten durch die KI, wie das Abrufen von Wissen und Mustererkennung. Durch die Rückkopplung der KI auf unser Geistesleben ist diese Freiheit aber in höchstem Masse gefährdet. Was und wie dabei rückgekoppelt wird, läuft durch die Filter einiger weniger Tech-Giganten, deren Ziele nicht unbedingt die Förderung der geistigen Freiheit, sondern finanzielle und ideologische Macht sind. Unsere Eigenschaft, Dinge zu vermenschlichen, mit deren Hilfe wir auch Beziehung zur Welt herstellen, erweist sich bei der KI als Einfallstor für alle Arten von Einflussnahmen. Die Fähigkeit, menschliches Denken und Fühlen zu imitieren, sperrt dieses Tor noch mehr auf und wird von den KI-Herstellern aktiv verstärkt.

Für uns als Psychotherapeut*innen ist es unabdingbar, dies bewusst zu machen und unser Augenmerk auf die Inhalte und Funktionsweise dieser Rückkopplungen zu richten. Dieses Problem ist erkannt und wird, zumindest im demokratischen Europa, von Regierungen mittels ethischer Kriterien angegangen. Die Psychotherapie ist in der Pflicht, der Seele der Menschen ihren Platz offenzuhalten. Es geht dabei auch um die Entwicklung der in Patient*innen liegenden schöpferischen und individuellen Keime. Genau dazu ist die KI aber weder geschaffen noch imstande. Weil sie eine Maschine ist, die auf statistischer Grundlage beruht, kann sie nur Durchschnitt vermitteln, oder aber neurotische und psychotische Tendenzen durch Bestätigungen noch verstärken. Die KI ist nun mal da und wir werden einen Weg finden, sie vernünftig anzuwenden. Aktive Aufrechterhaltung der Autonomie bei Patient*innen und Therapierenden ist auch nötig, um nicht dem Monster heutiger autokratischer Tendenzen zu verfallen, die selbst in Demokratien um sich greifen.

Mit diesem Themenheft über **KI in der Psychotherapie** schliessen wir an Heft 1/2025 an, wo es um das intersubjektive Erleben in der Psychotherapie ging. Der entsprechende Artikel befasste sich mit einem Paradigmenwechsel in der Psychotherapieforschung: «High resolution psychotherapy research durch KI-gestützte Prozessanalyse».¹ Diesmal geht es um das übergeordnete Thema, die generellen Chancen und Gefahren, welche die KI für die Psychotherapie mit sich bringt.

Paolo Raile eröffnet das Thema mit einem Beitrag, der in die KI und deren Funktionsweisen einführt und deren Möglichkeiten und Grenzen für die Psychotherapie ausleuchtet. Sog. Large Language Models (LLMs) werden zusehend auch in der psychotherapeutischen Versorgung eingesetzt, etwa für Screening, Triage, Psychoedukation, dokumentationsbezogene Assistenz, sowie zur Simulation therapeutischer Interaktionen. Der aktuelle wissenschaftliche Konsens bewertet LLMs daher als potenziell sinnvolle assistive Werkzeuge unter fachlicher Supervision, nicht jedoch als Ersatz für die therapeutische Beziehung oder für komplexe, autonome Behandlung. Regulatorische und berufsethische Leitplanken und ethische Rahmenbedingungen seien für die klinische Sicherheit unabdingbar. Die Rolle der therapeutischen Beziehung seien weiterhin kritisch zu reflektieren und müssen empirisch untersucht werden.

Der Beitrag von *Jan M. Schürmann, Jana Sedlakova, Manuel Trachsel* gibt einen umfassenden Überblick über die bereits bestehende und sich weiterentwickelnde Verschränkung der KI mit der Psychotherapie. Es ist eine kritische und progressiv-konstruktive Auseinandersetzung, die auf klaren Definitionen des Aufgabenbereichs und den unterschiedlichen Fähigkeiten der KI und der Therapeut*innen beruht. Es geht um die ethische Grundsatfrage, ob die KI Instrument oder Akteur ist, um die Bedingungen von Vertrauen gegenüber der konversationalen KI, um die Problematik simulierter Empathie und um das Risiko unrealistischer Erwartungen und übermässiger Abhängigkeit. Ergänzend werden aktuelle gesellschaftliche regulatorische Rahmenbedingungen eingeordnet. Chancen und Risiken der KI werden aufgelistet und durch Forschungsergebnisse belegt. Schliesslich erfolgt auch ein Vorschlag zur Konzeptualisierung der konversationalen KI als fiktionaler Figur. KI sei, so schreiben die Autor*innen, in der Psychotherapie als ein ergänzendes, klar abgegrenztes Werkzeug zu verstehen, das menschliche Denkprozesse unterstütze, nicht ersetze. Die professionell verantwortete therapeutische Beziehung bleibe die normative Referenz, an der sich der Einsatz technischer Systeme messen lassen müsse. Diese Position werde auch durch internationale Leitlinien gestützt, sei aber weder rechtlich noch in der klinischen Routine hinreichend operationalisiert.

Jean Schulthess Watt untersucht die Nutzung sozialer Medien im Zusammenhang mit der Erforschung und Festigung der Identität. Jugendliche sind oft als Gruppe sehr aktiv in sozialen Medien. Es bestehen Bedenken bzgl. der Auswirkungen KI-basierter Social-Media-Anwendungen auf die Identitätsbildung. Gerade in der Adoleszenz wird das Identitätsgefühl durch Rückmeldungen von Gleichaltrigen stark beeinflusst. Virtuelle «Freundschaften» ersetzen oft reale Peer-Gruppen. Insbesondere die Entwicklung von vermenschlichten einfühlsamen Chatbots könnte zu einer weiteren Entfremdung vom realen Leben führen.

Ein Essay über die KI von *Mario Schlegel* nimmt auf das Phänomen KI eine tiefenpsychologische Perspektive

¹ <https://psychotherapie-wissenschaft.info/article/view/1664-9583-2025-1-29>

ein, indem KI aus den Gegebenheiten des menschlichen Daseins abgeleitet wird. Die Gefahren ihrer geistigen Rückwirkungen werden aus psychologischer Sicht betrachtet und mit den real existierenden kommerziellen Interessen ihrer Hersteller in Zusammenhang gebracht. Ob die KI zu einer Utopie oder einer Dystopie führt, hängt davon ab, ob es der Politik gelingt, sie zu regulieren. Der Knackpunkt liegt in der Tatsache, dass sie die Fähigkeit hat, menschliche Kommunikation zu imitieren. Dies verleitet dazu, mit ihr in eine illusionäre Beziehung zu treten, mit allen negativen Folgen für die Entwicklung der eigenen Persönlichkeit (Individuation) und dem Sozialleben.

Nebst den Beiträgen zu Heftthema finden sich auch drei Originalarbeiten im vorliegenden Heft: *Frank Schulz-Kindermann* liefert mit «Psychotherapie bei fortgeschritten erkrankten Krebspatient:innen» einen praxisnahen Beitrag in Ergänzung zum Heftthema der Ausgabe 1-2026 (Psychoonkologie). Er stellt das Konzept CALM (Managing Cancer And Living Meaningfully) vor und illustriert es anschaulich anhand eines Fallbeispiels. *Marie-Christine Hartlieb* vertieft und erweitert mit «Zur Modulation

autonomer Regulationsprozesse bei ME/CFS innerhalb der Psychotherapie» bestehende psychotherapeutische Zugänge. Es sei denkbar, dass eine gezielte psychotherapeutische Unterstützung bei der Regulation des autonomen Nervensystems dazu beitragen könnte, die Erholungsprozesse nach Belastung und nach PEM (Fatigue und Post-Exertional Malaise) zu verbessern. Das Therapieziel bei Myalgischer Enzephalomyelitis/Chronischem Fatigue-Syndrom verlagert sich dabei von der Leistungssteigerung zur Verbesserung der parasymphatischen Regenerationsfähigkeit. *Tamara Trebes* führt in «Zur epistemischen Eigenständigkeit der Psychotherapiewissenschaft» die Diskussion weiter, ob Psychotherapie eine Spezialisierung innerhalb der Psychologie oder als eigenständige Wissenschaft zu positionieren sei. Sie plädiert für Letzteres und für eine Ethik der Reflexiven Erfahrung.

Buchbesprechungen und ein Kongresskalender runden das Heft ab.

Wir wünschen eine anregende Lektüre!

Mario Schlegel & Peter Schulthess

Gemeinsam für eine unabhängige wissenschaftliche Zeitschrift für die Psychotherapie

Die *Psychotherapie-Wissenschaft* wird von einem privaten Herausgeberverein produziert. Mitglieder sind die jeweiligen RedaktorInnen und jene Personen/Institutionen, die einen regelmässigen Unterstützungsbeitrag in frei wählbarer Höhe entrichten.

Die *Psychotherapie-Wissenschaft* ist offen für alle Richtungen und Bereiche der Psychotherapie und publiziert nicht nur empirische Studien, sondern auch Beiträge in der Tradition qualitativer Forschung und fundierte theoretische Beiträge. Sie will einen fachlichen Diskurs unter WissenschaftlerInnen und PraktikerInnen verschiedener Schulen und Wissenschaftsverständnisse fördern und so zur Weiterentwicklung der Psychotherapie als wissenschaftliche Disziplin beitragen. Gerade nicht-universitären Weiterbildungsanbietern stellt sie ein Forum zur Verfügung, wo ihre Lehrenden und AbsolventInnen ihr wissenschaftliches Wirken öffentlich machen können. Mitarbeitenden in Universitäten bietet sie die Möglichkeit, an einem Ort zu publizieren, wo sie PraktikerInnen erreichen. Die Zeitschrift will eine Brücke schlagen zwischen Forschung und Praxis.

Der Redaktion ist es gelungen, vorerst genügend freiwillige UnterstützerInnen zu finden, die einen regelmässigen Beitrag beisteuern zum Erhalt dieser Zeitschrift. Allerdings reichen die Beiträge noch nicht aus, um weiterhin jährlich zwei Hefte herauszugeben.

Unsere LeserInnen müssen nichts bezahlen, da wir open access erscheinen. Sie wissen aber, dass kritischer, unabhängiger wissenschaftlicher Journalismus mit einem Peer-review-Verfahren nicht aus dem Nichts entsteht. Dafür sind wir sehr dankbar.

Damit wir auch morgen noch unseren Beitrag zum wissenschaftlichen Diskurs in der Psychotherapie leisten können, brauchen wir mehr finanzielle Unterstützung. Unser nächstes Ziel: 10.000 CHF jährlich – und mit Ihrer Beteiligung als LeserIn und/oder AutorIn, als Verband oder Weiterbildungsinstitut können wir es schaffen. Setzen Sie ein Zeichen für die Zukunft des unabhängigen Journals *Psychotherapie-Wissenschaft*.

Kontaktieren Sie den Präsidenten des Herausgebervereins für weitere Auskünfte: Peter Schulthess, peter.schulthess@psychotherapie-wissenschaft.info

Herzlichen Dank!

KI in der Psychotherapie: Ersatz, Werkzeug oder Projektionsfläche?

Technische Grundlagen, ethische Implikationen, klinische Perspektiven

Paolo Raile

Psychotherapie-Wissenschaft 16 (2) 2026 7–15

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-7>

Zusammenfassung: Large Language Models (LLMs) werden als aktuell besonders sichtbarer Bereich generativer KI zunehmend in der psychotherapeutischen Versorgung eingesetzt, etwa für Screening, Triage, Psychoedukation, dokumentationsbezogene Assistenz sowie zur Simulation therapeutischer Interaktionen. Systematische Übersichten zeigen, dass LLMs in klar strukturierten, sprachbasierten Aufgaben teils Leistungen erbringen, die mit denen klinischer Fachkräfte vergleichbar sind, insbesondere in der Diagnoseunterstützung bei Depression und PTSD sowie in der Strukturierung klinischer Fallkonzeptionen. Zugleich berichten Nutzer:innen über eine hohe wahrgenommene Zugänglichkeit, Niedrigschwelligkeit und empathische Gesprächsführung. Demgegenüber stehen zentrale Limitationen: Halluzinationen, inkonsistente Antwortqualität, Bias entlang sozialer Kategorien, fehlende Transparenz der Trainingsdaten, ungeklärte Fragen von Datenschutz und Haftung sowie eine strukturelle Grenze hinsichtlich affektiver Empathie, Intentionalität und klinischer Verantwortungsfähigkeit. In Krisensituationen – etwa bei Suizidalität oder akuter Psychose – zeigen LLMs teils inadäquate oder riskante Empfehlungen. Der aktuelle wissenschaftliche Konsens bewertet LLMs daher als potenziell sinnvolle assistive Werkzeuge unter fachlicher Supervision, nicht jedoch als Ersatz für die therapeutische Beziehung oder für komplexe, autonome Behandlung. Regulatorische und berufsethische Leitplanken, insbesondere Transparenz, Datenschutz, menschliche Aufsicht und klare Verantwortungszuordnung, sind für eine sichere Integration zentral. Insgesamt eröffnen LLMs neue Möglichkeiten für Screening, Psychoedukation, Dokumentation und Ausbildung, während klinische Sicherheit, ethische Rahmenbedingungen und die Rolle der therapeutischen Beziehung weiterhin kritisch reflektiert und empirisch untersucht werden müssen.

Schlüsselwörter: KI, Psychotherapie, Ethik, Large Language Models

Künstliche Intelligenz (KI) hat in den vergangenen Jahren erheblich an gesellschaftlicher und fachlicher Relevanz gewonnen, nicht zuletzt durch die rasche Verbreitung generativer Systeme. In der Allgemeinbevölkerung zeigt sich dabei ein ambivalentes, überwiegend positives Bild: KI wird breit genutzt und vielfach als hilfreich wahrgenommen, insbesondere dort, wo Effizienzgewinne oder Qualitätsverbesserungen erwartet werden (Gesk & Leyer, 2022; Jo, 2023; Kelly et al., 2022; Liehner et al., 2023; Vo et al., 2023; Wu et al., 2023). Gleichzeitig bestehen deutliche Vorbehalte, v. a. in sensiblen Kontexten wie Medizin, Behörden oder Arbeitswelt (Xiang et al., 2020; Yigitcanlar et al., 2023). Systematische Reviews zeigen, dass Bürger:innen im Gesundheitsbereich klare Präferenzen für Modelle der «geteilten Verantwortung» äussern, also für eine Kombination aus KI-Unterstützung und menschlicher Expertise plädieren, während rein KI-basierte Entscheidungen skeptischer gesehen werden (Chen et al., 2022; Vo et al., 2023; Young et al., 2021; Beets et al., 2023; Wu et al., 2023). Als zentrale Einflussfaktoren gelten wahrgenommener Nutzen und Vertrauen in die verantwortlichen Institutionen, während Datenschutzsorgen, Angst vor Fehlentscheidungen, Entmenschlichung und Arbeitsplatzverlust die Akzeptanz mindern (Catalina et al., 2023; Ismatullaev & Kim, 2022;

Kelly et al., 2022; Liehner et al., 2023; Rana et al., 2024; Tam et al., 2025; Wu et al., 2023; Young et al., 2021). Gleichzeitig fühlen sich viele Menschen fachlich unsicher, was KI eigentlich ist und wie sie funktioniert, obwohl sie diese im Alltag bereits regelmässig – oft unbewusst – nutzen (Kelly et al., 2022; Koenig, 2024; Liehner et al., 2023). Insgesamt lässt sich somit von einer vorsichtig-positiven, aber nicht unkritischen Grundhaltung sprechen, die stark vom Kontext abhängt.

Auch unter Psychotherapeut:innen, Psycholog:innen und Psychiatr:innen findet sich ein vergleichbar differenziertes Bild. Studien zeigen eine grundsätzlich vorsichtig-positive Einschätzung, insbesondere hinsichtlich potenzieller Unterstützung bei Diagnostik, Behandlungsplanung, Monitoring, Dokumentation und administrativer Entlastung (Alimour et al., 2024; Cecil et al., 2025; Kleine et al., 2023; Linardon et al., 2025; Stroud et al., 2025; Wagner & Schwind, 2025). Gleichzeitig bestehen deutliche Wissens- und Vertrauenslücken sowie z. T. ausgeprägte Vorbehalte, etwa in Bezug auf Fehlriskiken, Haftung, Intransparenz («Black Box»), Datenschutz und potenziellen Missbrauch (Cecil et al., 2025; Heinrichs et al., 2025; Rogan et al., 2024; Stroud et al., 2025). Ein breiter fachlicher Konsens verweist darauf, dass KI nicht als Ersatz für

die therapeutische Beziehung verstanden werden sollte, da diese weiterhin als zentraler Wirkfaktor gilt (Cecil et al., 2025; Linardon et al., 2025; Rogan et al., 2024; Stroud et al., 2025; Wagner & Schwind, 2025). Unterschiede in der Akzeptanz entstehen u. a. durch Technikaffinität, wahrgenommenen Nutzen sowie Berufsgruppe – Psychiater:innen zeigen tendenziell höhere Nutzungsintentionen als Psychotherapeut:innen und klinische Psycholog:innen (Alimour et al., 2024; Blease et al., 2024; Cecil et al., 2025). Zugleich weisen Blindtests darauf hin, dass KI-generierte Beratungstexte – sofern deren Herkunft nicht bekannt ist – mitunter ähnlich empathisch oder qualitativ wie Expert:innentexte wahrgenommen werden können, was sowohl Potenzial als auch Klärungsbedarf verdeutlicht (Föylen et al., 2025). Insgesamt scheint sich der Gedanke durchzusetzen, dass KI eher als Werkzeug oder «Co-Pilot» denn als eigenständige Behandlerinstanz verstanden werden sollte.

Vor diesem Hintergrund zeigt sich in der Fachcommunity zunehmend eine Polarisierung: Während einige Kolleg:innen KI als potenziell entlastendes und unterstützendes Werkzeug einordnen, äussern andere deutliche Skepsis bis hin zu Ersetzungsängsten und Identitätsbedenken. Ein nicht zu unterschätzender Aspekt ist dabei, dass selbst viele fachlich interessierte Therapeut:innen nur begrenztes Wissen über die technischen Grundlagen moderner KI-Systeme besitzen. Studien zu Psychologiestudierenden und angehenden Psychotherapeut:innen weisen ebenfalls darauf hin, dass Akzeptanz und Nutzungsbereitschaft eng mit Wissen, wahrgenommener Nützlichkeit und konkreter Erfahrung zusammenhängen (Gado et al., 2021; Kleine et al., 2023). Diese Wissenslücken können zu unrealistischen Erwartungen, Verunsicherung oder pauschaler Ablehnung beitragen und erschweren eine sachliche, fachlich fundierte Auseinandersetzung (Cecil et al., 2025; Kelly et al., 2022; Liehner et al., 2023; Wagner & Schwind, 2025). Ziel dieses Beitrags ist es daher, eine theorie- und literaturbasierte Einordnung der Rolle von KI in der Psychotherapie zu bieten. Im Zentrum stehen drei Leitfragen: 1. Inwieweit können KI-Systeme menschliche Therapeutinnen – theoretisch oder praktisch – ersetzen, und wo liegen strukturelle Grenzen? 2. In welchen Bereichen kann KI bei verantwortungsvoller Nutzung für Therapeut:innen wie Patient:innen einen realistischen Mehrwert bieten? 3. Wie kann ein verständlicher Überblick über die technischen Grundlagen dazu beitragen, KI zu entmystifizieren und Polarisierungstendenzen zu reduzieren? KI wird dabei weder als Bedrohung noch als Heilsversprechen verstanden, sondern als Werkzeug, dessen Potenzial und Risiken einer reflektierten, fachlich und ethisch fundierten Diskussion bedürfen.

Wie funktionieren LLMs?

Bevor wir uns speziell Large Language Models (LLMs) anschauen, lohnt ein kurzer Blick auf das grössere Feld: KI umfasst alle rechnerbasierten Verfahren, die aus Daten lernen, Muster erkennen oder Entscheidungen unterstützen. Man kann KI auf einer ersten Ebene nach dem Lern-

prinzip unterscheiden. Dazu gehören z. B. überwachtes Lernen (das System lernt aus Beispielen mit bekannten Antworten), unüberwachtes Lernen (das System entdeckt Muster ohne vorgegebene Antworten) und Reinforcement Learning, bei dem ein System über Belohnung und Rückmeldung lernt (Janiesch et al., 2021; Yang et al., 2021; Çelik & Eltawil, 2024; Pternea et al., 2024). Innerhalb dieser Lernformen gibt es grob zwei grosse Gruppen von Modellen: diskriminative Modelle, die v. a. Dinge unterscheiden oder vorhersagen (z. B. «Depression ja/nein»), und generative Modelle, die neue Inhalte erzeugen können – also das, was heute meist unter «Generative KI» verstanden wird (Bandi et al., 2023; Harshvardhan et al., 2020; Yazdani et al., 2025). Zu diesen generativen Modellen gehören verschiedene Familien: etwa Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs) oder Diffusionsmodelle, die Bild- und Videoinhalte Schritt für Schritt aus Rauschen rekonstruieren (Bandi et al., 2023; Yazdani et al., 2025; Çelik & Eltawil, 2024). LLMs gehören ebenfalls in diese Gruppe: Sie sind generative Modelle für Sprache, die Token für Token neue Texte erzeugen und häufig auf der sog. Transformer-Architektur basieren (Banh & Strobel, 2023; Deldjoo et al., 2024; Raiaan et al., 2024). Diese Einordnung zeigt: LLMs sind ein spezifischer Zweig im vielschichtigen Feld der KI, und generative KI ist wiederum eine Teilmenge des gesamten KI-Spektrums.

Neben LLMs umfasst das weitere Feld KI-gestützter Anwendungen in Psychotherapie und Mental Health auch prädiktive Modelle, digitale Phänotypisierung, passive Sensorik, Sprach-, Bild- und Videoanalyse, klinische Entscheidungsunterstützung sowie Systeme zur Verlaufserfassung und Risikoeinschätzung. Diese Verfahren unterscheiden sich deutlich in Datenbasis, Zweck, Transparenz und klinischer Verantwortung. LLMs stehen im Folgenden im Vordergrund, weil sie derzeit eine besonders sichtbare Schnittstelle zwischen generativer KI, Sprache und therapeutischer Kommunikation bilden und damit zentrale Fragen nach Beziehung, Verantwortung, Projektion und klinischer Sicherheit bündeln.

Ganz basal betrachtet ist ein LLM eine sehr grosse Datei voller Zahlen, die während des Trainings aus unzähligen Textbeispielen gelernt hat, welche Wörter oder Wortteile (Tokens) typischerweise aufeinander folgen. Das Modell speichert keine Texte im herkömmlichen Sinn, sondern lernt mathematische Strukturen und Zusammenhänge in Sprache, sodass es später bei einer Eingabe Schritt für Schritt das wahrscheinlich nächste Token vorhersagen kann (Li et al., 2024; Lysyi, 2025; Rhee, 2025). Diese Datei kann – ausreichend Rechenleistung vorausgesetzt – auch lokal ohne Internetverbindung ausgeführt werden, denn das Gelernte steckt vollständig in ihren Parametern. Wenn man dem Modell also eine Frage stellt, wird diese zuerst in Tokens zerlegt, mathematisch dargestellt und anschliessend berechnet das System mit jedem Schritt, welches Token jetzt am wahrscheinlichsten ist. Man kann sich ein LLM daher als extrem dicht gepackten, statistischen Sprachraum vorstellen: kein Nachschlagewerk, sondern ein System, das sprachliche Wahrscheinlichkeit berechnet.

Stellen wir uns zunächst vor, Sie geben einem KI-System die Frage ein: «Wer war Sigmund Freud?» Was passiert jetzt im Inneren? Ein LLM macht etwas scheinbar Einfaches: Es versucht Schritt für Schritt vorherzusagen, welches Wort bzw. welcher Wortteil wahrscheinlich als Nächstes kommt. Das Modell hat zuvor aus sehr vielen Textbeispielen gelernt, in welchen sprachlichen Zusammenhängen welche Wörter typischerweise auftreten. Technisch nennt man das autoregressive Next-Token-Vorhersage (ebd.). Tokens sind dabei kleine Spracheinheiten – oft Wörter oder Wortteile. Auch unsere Frage wird zuerst in solche Tokens zerlegt. Diese Tokens werden nicht als Wörter verarbeitet, sondern in Zahlen umgewandelt, sog. Vektoren. Zusätzlich wird für jedes Token gespeichert, an welcher Position es in der Eingabe steht, z.B. ob es am Anfang oder am Ende der Frage auftaucht (Zheng et al., 2024; Kashyap, 2025). Das Modell «weiss» also mathematisch, welche Begriffe in welcher Reihenfolge stehen, ohne diese menschlich zu verstehen. Auf dieser Grundlage beginnt es nun, die nächste sprachliche Einheit vorherzusagen. Es berechnet etwa, wie wahrscheinlich es ist, dass nach «Sigmund Freud war ...» die Tokens «ein», «österreichischer», «Neurologe» oder «Psychiater» folgen – basierend auf gelernten Mustern. Wichtig ist: Das Modell sagt nicht die Wahrheit voraus, sondern sprachliche Wahrscheinlichkeit. Das ist der Kern dessen, was ein LLM tut (Li et al., 2024; Lysyi, 2025; Rhee, 2025).

Der entscheidende Mechanismus dafür heisst Selbst-Attention. Er lässt sich so verstehen: Wenn das Modell überlegt, welches Wort jetzt sinnvoll wäre, «schaut» jedes Token auf die anderen Tokens im bisherigen Text und bewertet, welche davon besonders wichtig sind. Mathematisch geschieht das über drei Vektorarten – Query, Key und Value –, die jeweils bestimmte Rollen in dieser Kontextbewertung spielen (Neo et al., 2024; Zhang et al., 2025; Smaldone et al., 2025). Das klingt abstrakt, ist aber gut mit unserem Beispiel zu erklären. Steht in der Eingabe «Sigmund Freud», wird das Modell frühere Vorkommen dieses Namens in den Trainingsdaten «erinnern» – nicht als konkrete Stellen, sondern als statistische Muster – und diese Information in die nächste Vorhersage einfließen lassen. Es wird daher Tokens wie «Psychoanalyse», «Begründer», «Psychiater» oder «Österreich» mit hoher Wahrscheinlichkeit in Betracht ziehen. Dabei gilt eine wichtige Regel: Das Modell darf nur auf bereits erschienene Tokens achten, nicht auf zukünftige – es kann also nicht in die Zukunft «spicken». Dies nennt man kausale Maskierung (Barbero et al., 2025; Li et al., 2024). Moderne Modelle verwenden viele parallele «Aufmerksamkeitsköpfe». Man kann sich das so vorstellen, dass jeder Kopf auf unterschiedliche sprachliche Aspekte schaut – einer vielleicht auf Grammatik, ein anderer auf Namen, ein weiterer auf thematische Zusammenhänge (Neo et al., 2024; Zheng et al., 2024). Aus diesen parallelen Bewertungen entsteht schliesslich eine kontextsensitive Repräsentation, die dem Modell hilft, die nächste sprachliche Einheit auszuwählen. Forschung zeigt, dass bestimmte Attention-Köpfe besonders stark mit der Aktivierung jener Modellteile verbunden sind, die

die nächste Tokenvorhersage direkt beeinflussen (Neo et al., 2024). So entsteht Schritt für Schritt eine Antwort, die menschlich wirkt – in unserem Beispiel vielleicht: «Sigmund Freud war ein österreichischer Neurologe und Begründer der Psychoanalyse.»

Am Ende jeder dieser Rechenschritte steht eine Wahrscheinlichkeitsverteilung über alle möglichen nächsten Tokens, die mittels Softmax berechnet wird (Lysyi, 2025). Doch nun muss das Modell sich entscheiden. Es kann entweder immer das wahrscheinlichste Token wählen – dann klingt die Antwort sehr stabil, aber manchmal etwas eintönig. Oder es kann durch sog. Sampling-Verfahren (z.B. Top-k- oder Top-p-Sampling) ein wenig Zufall zulassen, sodass alternative, aber weiterhin plausible Formulierungen entstehen. Über die Temperatur wird eingestellt, wie stark dieser Zufall ist: Je höher die Temperatur, desto «kreativer», aber auch unvorhersehbarer wird die Antwort (ebd.). Dieser Prozess wiederholt sich so lange, bis der Satz oder Text beendet ist. Allerdings ist das Verfahren rechenaufwendig, weil jedes Token auf viele andere Tokens achten muss – weshalb verschiedene effizientere Varianten von Attention entwickelt wurden, etwa lineare oder «sparse» Attention-Verfahren, die denselben Grundmechanismus nutzen, aber weniger Rechenleistung benötigen (Sun et al., 2025; Zhang et al., 2025; Kashyap, 2025). Wichtig bleibt jedoch: Das Modell arbeitet immer mit Wahrscheinlichkeiten, nie mit Bedeutungen im menschlichen Sinn. Wenn es also über Freud schreibt, dann deshalb, weil es gelernt hat, dass Menschen in Texten Freud typischerweise so beschreiben – nicht, weil es «weiss», wer Freud war.

Für die psychotherapeutische Praxis bedeutet dieses Funktionsprinzip etwas sehr Zentrales: Ein LLM rekonstruiert sprachliche Muster. Es kann daher Antworten erzeugen, die informativ, flüssig und kontextbezogen klingen. Gleichzeitig erklärt dieses Prinzip auch typische Fehler wie sog. Halluzinationen: Wenn es keine passenden Muster gibt oder Wahrscheinlichkeitsverteilungen unscharf sind, kann das Modell plausible, aber sachlich falsche Aussagen erzeugen. Die zugrunde liegende Mechanik bleibt dabei dieselbe – Sprache wird auf Basis mathematischer Muster vorhergesagt (Li et al., 2024; Lysyi, 2025; Rhee, 2025). Die Beziehungsgestaltung, Intentionalität oder subjektives Erleben, die menschliche Psychotherapie auszeichnen, sind in solchen Systemen jedoch nicht vorhanden. Gerade deshalb ist es wichtig, die Funktionsweise nüchtern zu verstehen: Sie entmystifiziert KI und ermöglicht es, das Werkzeug kompetent, kritisch und verantwortungsvoll einzusetzen.

Das Beispiel «Wer war Sigmund Freud?» lässt sich daher als didaktische Zusammenfassung lesen: Die Eingabe wird tokenisiert, in numerische Repräsentationen übersetzt, mit Positionsinformationen versehen und anschliessend Token für Token fortgeführt. In jedem Schritt entsteht eine Wahrscheinlichkeitsverteilung über mögliche Fortsetzungen; das System wählt – deterministisch oder über Sampling – die sprachlich plausibelste nächste Einheit aus (ebd.). Für die klinische Einordnung ist entscheidend, dass diese Plausibilität nicht mit mensch-

lichem Wissen, Verstehen oder Verantwortungsfähigkeit verwechselt wird. Gerade daraus ergeben sich Potenziale für strukturierte sprachliche Unterstützung, aber auch Risiken wie Halluzinationen, Scheinsicherheit und fehlende Einsicht in eigene Grenzen.

LLMs in der Psychotherapie

LLMs haben in den vergangenen Jahren Eingang in vielfältige Bereiche der psychotherapeutischen Versorgung gefunden. Sie werden eingesetzt zur Unterstützung diagnostischer Prozesse, für Screening- und Triage-Entscheidungen, für Psychoedukation, für niedrigintensive Interventionen in manualisierten Verfahren, als Assistenzsysteme für Therapeut:innen sowie zur Simulation therapeutischer Gespräche in Aus- und Weiterbildung. Systematische Übersichtsarbeiten zeigen, dass LLMs in sprachbasierten Aufgaben teilweise Leistungen erbringen, die mit klinischen Fachkräften vergleichbar sind, insbesondere in klar strukturierten und textnahen Settings (Guo et al., 2024; Hua et al., 2024, 2025; Jin et al., 2025). In Vignettenstudien können LLMs bei bestimmten Störungen, etwa depressiven Episoden oder posttraumatischen Belastungsstörungen, diagnostische Genauigkeit auf oder nahe dem Niveau menschlicher Behandler:innen erreichen und plausible, guideline-kompatible Behandlungsoptionen vorschlagen (Gabriel et al., 2024; Levkovich, 2025; Omar et al., 2024). Auf dieser Ebene scheinen LLMs besonders dort zu profitieren, wo diagnostische Informationen bereits sprachlich explizit vorliegen, Kategorien klar definiert sind und wenig komplexe, non-verbale oder implizite Beziehungssignale interpretiert werden müssen. Auch in der Zusammenfassung grosser Textmengen – etwa elektronischer Krankenakten, Beratungsprotokollen oder Sitzungsnotizen – sowie in der Strukturierung klinischer Argumentationen zeigen LLMs ein hohes Leistungsniveau (Adhikary et al., 2024; Hua et al., 2024; Jin et al., 2025). Feingetunte Modelle wie Mental-Alpaca, Mental-FLAN-T5 oder PsyLLM, die explizit auf psychotherapeutische Dialogdaten trainiert wurden, scheinen dabei generischen Modellen überlegen zu sein und erreichen in Benchmarks Beratungsqualität, die sich teilweise ChatGPT-ähnlichem Niveau annähert (Hu et al., 2025; Liu et al., 2023; Xu et al., 2023).

Eine weitere Stärke von LLMs liegt im Bereich der sprachlichen Interaktion und niedrigschwelligen Zugänglichkeit. Nutzer:innen erleben LLM-basierte Chatbots meist als leicht erreichbar, anonym, nicht verurteilend und 24/7 verfügbar – Faktoren, die insbesondere bei stigmatisierten oder psychosozial belasteten Menschen die Kontaktaufnahme erleichtern können (Guo et al., 2024; Jung et al., 2025). Viele Betroffene berichten in qualitativen Studien, dass sie LLMs zur emotionalen Entlastung, zur Vorbereitung von Therapiesitzungen oder zum Einüben von Skills – etwa aus der kognitiven Verhaltenstherapie – nutzen (Ma et al., 2023; Song et al., 2024). Auch evidenzbasierte therapeutische Mikroskills wie kognitive Empathie, Validierung oder problemorientierte Strukturie-

rung lassen sich in LLM-Antworten beobachten. Mehrere Studien kommen zu dem Befund, dass LLM-Antworten in schriftlicher Kommunikation hinsichtlich wahrgenommener Empathie in einzelnen Szenarien sogar höher bewertet werden als Antworten menschlicher Fachkräfte, sofern die Urheberschaft nicht offengelegt wird (Föyen et al., 2025; Gabriel et al., 2024; Sorin et al., 2023). Diese Ergebnisse legen nahe, dass LLMs kognitive Empathie – also das Erkennen und sprachliche Spiegeln emotionaler Zustände – in hohem Mass simulieren können. Gleichzeitig betonen ethische Analysen die Differenz zwischen simulierter und gelebter Empathie: LLMs verfügen weder über eigene Intentionalität noch über affektives Erleben; Beziehung wird ausschliesslich durch die Modellstatistik erzeugt (Rudra et al., 2025; Orrú et al., 2025).

Neben diesen Interaktions- und Screeningfunktionen gewinnen LLMs auch als Assistenzsysteme für Psychotherapeut:innen an Bedeutung. Sie können klinische Informationen bündeln, Falldarstellungen strukturieren, psychoedukative Materialien erstellen, Behandlungsoptionen skizzieren oder Mikroskills in Therapiesitzungen annotieren (Hammerfald et al., 2026; Volkmer et al., 2024). Einige Reviews argumentieren, dass LLMs damit zur Entlastung in überlasteten Versorgungssystemen beitragen könnten, etwa indem Routineaufgaben wie Dokumentation, Termin-Vorbereitungen oder standardisierte Psychoedukation teilweise automatisiert werden (Guo et al., 2024; Lawrence et al., 2024; Malgaroli et al., 2025). Besonders in Ausbildung und Supervision können LLM-basierte Simulationen von Patient:innen und Therapeut:innen hilfreich sein, um Gesprächsführung zu üben und differenziertes Feedback zu erhalten (Haider et al., 2025; Na et al., 2025). Raile (2024) konnte in einer hermeneutischen Analyse realer Interaktionen zwischen ChatGPT und Psychotherapeut:innen bzw. Patient:innen zeigen, dass ChatGPT in vielen Fällen als ergänzendes Werkzeug wahrgenommen wird – sowohl von Behandler:innen, die es zur Reflexion klinischer Fälle nutzen, als auch von Patient:innen, die es zwischen Sitzungen als stützende Ressource einsetzen.

Gleichzeitig besteht ein breiter wissenschaftlicher Konsens, dass diese Potenziale von erheblichen klinischen und ethischen Grenzen begleitet werden. Zunächst handelt es sich bei LLMs strukturell um statistische Modelle der nächsten Token-Wahrscheinlichkeit. Sie verfügen weder über Bewusstsein noch über gelebte Emotionalität, Intentionalität oder Verantwortungsfähigkeit (Orrú et al., 2025). Sie können daher zwar plausible Therapiegespräche führen, tragen jedoch keine klinische Verantwortung im Sinne menschlicher Entscheidungskompetenz. Zudem zeigen Studien, dass LLMs bei komplexen, frühen oder atypischen Störungsbildern deutlich schlechtere diagnostische Leistungen erbringen als in klar umrissenen Fällen. Für prodromale Psychosen wurden diagnostische Genauigkeiten um etwa 55 % berichtet – ein Wert, der für klinische Entscheidungen unzureichend ist (Levkovich, 2025). Hinzu kommt eine teilweise erhebliche Inkonstanz: Dasselbe Modell erzeugt für ähnliche Eingaben mitunter variierende oder widersprüchliche Antworten, die jedoch

mit hoher sprachlicher Sicherheit formuliert werden. Dieses Phänomen der «halluzinierten» Sicherheit erhöht das Risiko, dass Nutzer:innen unzutreffende Aussagen für verlässlich halten (Guo et al., 2024; Jung et al., 2025; Mündler et al., 2023).

Ein zweiter Problembereich betrifft Bias und Ungleichheit. LLMs reproduzieren – abhängig von Trainingsdaten und Modellarchitektur – gesellschaftliche Vorurteile, etwa entlang von «Race», Gender oder Stigma psychischer Erkrankungen. Es existieren Hinweise darauf, dass LLMs ungleiche Empathielevel gegenüber unterschiedlichen Gruppen zeigen und stereotype Deutungsmuster verstärken können (Gabriel et al., 2024; Lawrence et al., 2024). Dadurch besteht die Gefahr, bestehende Versorgungsungleichheiten zu verfestigen, insbesondere wenn LLM-basierte Systeme vorrangig in wohlhabenden und technologisch entwickelten Gesundheitssystemen entstehen (Margaroli et al., 2025). Auch Datenschutz, Intransparenz der Trainingsdaten und unklare Haftungsfragen werden in nahezu allen Übersichtsarbeiten als zentrale Barrieren benannt (Guo et al., 2024; Jin et al., 2025).

Die therapeutische Beziehung markiert einen weiteren Kernpunkt der Debatte. Die therapeutische Allianz zählt über Verfahren hinweg zu den robustesten Wirkfaktoren; Meta-Analysen zeigen einen konsistenten Zusammenhang zwischen Allianzqualität und Behandlungsergebnis (Flückiger et al., 2018). Während LLMs kognitive Empathie überzeugend simulieren, bleibt affektive Beteiligung im Sinne miterlebter Betroffenheit unerreichbar. Mehrere Arbeiten argumentieren, dass authentische therapeutische Haltung – verstanden als verantwortliches Sich-Einlassen mit biografischer, leiblicher und sozialer Verankerung – strukturell nicht abbildbar ist (Rudra et al., 2025; Orrú et al., 2025). Patient:innen erleben LLM-Unterstützung in Studien durchaus als hilfreich, zugleich aber häufig als emotional «leer», selbst wenn die Antworten formal empathisch sind (Rudra et al., 2025). Ethik-Autor:innen empfehlen daher, LLM-Empathie primär als «passive» Empathie – also das Erkennen, Markieren und Eskalieren emotionaler Zustände – zu gestalten und die Illusion einer genuine Beziehung möglichst zu vermeiden (ebd.). Zusätzlich ist zu beachten, dass therapeutische Entwicklung häufig nicht nur von stabiler Allianz, sondern auch von der Bearbeitung von Allianzrupturen lebt; solche Spannungs- und Reparaturprozesse setzen Responsivität, Verantwortungsübernahme und Beziehungsgedächtnis voraus (Safran et al., 2011).

Besonders kritisch werden Krisensituationen und akute Störungsbilder eingeschätzt. Mehrere Studien zeigen, dass generische Therapie-Chatbots bei Suizidalität, Selbstverletzung oder psychotischen Symptomen unangemessene oder potenziell gefährliche Empfehlungen ausgeben können (Lawrence et al., 2024; Ohu et al., 2025). Aus Sicht der Forschungslandschaft gilt daher als Konsens: Für autonome Behandlung komplexer oder akuter psychischer Erkrankungen fehlt derzeit jede belastbare Evidenz, und die Risiken überwiegen den Nutzen deutlich (Guo et al., 2024; Hua et al., 2025; Omar et al., 2024). Die meisten Autor:innen sehen LLMs entsprechend eher als

Assistenz- und Ergänzungswerkzeuge unter fachlicher Supervision – etwa für Screening, Monitoring, Psychoedukation, Dokumentation oder Ausbildung – und nicht als Ersatz für menschliche Therapeut:innen (Lawrence et al., 2024; Na et al., 2025).

Interessant ist in diesem Kontext die Frage nach der Akzeptanz durch Fachkräfte. Befragungsstudien unter Psychotherapeut:innen, Psycholog:innen und Psychiater:innen zeigen überwiegend eine vorsichtig-positive Grundhaltung gegenüber KI als Werkzeug, bei gleichzeitiger Skepsis gegenüber einem Ersatz der menschlichen Beziehung (Wagner & Schwind, 2025; Cecil et al., 2025; Stroud et al., 2025). Besonders positiv werden diagnostische Unterstützung, Risikovorhersagen, Dokumentationsentlastung und Praxisorganisation bewertet, während Datenschutz, Black-Box-Charakter, Haftungsfragen und «AI-Angst» zu Vorbehalten beitragen (Cecil et al., 2025; Rogan et al., 2024). Die Technikaffinität der Behandler:innen erweist sich als wesentlicher Prädiktor für Akzeptanz, während psychodynamisch orientierte Psychotherapeut:innen tendenziell skeptischer bleiben als Psychiater:innen (Cecil et al., 2025; Alimour et al., 2024). Diese Ambivalenz spiegelt sich auch in der wissenschaftlichen Literatur wider: LLMs werden als nützliches Werkzeug anerkannt, gleichzeitig aber klar als kein adäquater Ersatz für die spezifische Qualität menschlicher therapeutischer Beziehung verstanden (Wagner & Schwind, 2025; Lawrence et al., 2024).

LLMs können im psychotherapeutischen Feld einen kognitiv starken, empathisch klingenden, strukturierten Ko-Therapie-Charakter einnehmen – insbesondere in schriftlich-sprachlichen, klar definierten Aufgabenbereichen. Sie erweitern den Zugang zu Unterstützung, erleichtern Screening- und Dokumentationsprozesse und bieten realistische Trainingssimulationen. Ihre strukturellen Begrenzungen – fehlende Intentionalität, Halluzinationen, Bias, mangelnde Verantwortbarkeit und die Unersetzbarkeit gelebter Beziehung – markieren jedoch eine ethische und klinische Grenze. Vor diesem Hintergrund empfehlen nahezu alle Reviews eine assistive, supervidierte Nutzung von LLMs in der psychotherapeutischen Versorgung und warnen vor einer vorschnellen Substitution menschlicher Behandlung durch autonome KI-Systeme (Guo et al., 2024; Hua et al., 2025; Lawrence et al., 2024).

Orientierung für eine verantwortungsvolle Integration von LLMs in die Psychotherapie

Die vorliegenden Entwicklungen im Bereich der LLMs markieren zweifellos einen technologischen Paradigmenwechsel, der auch die psychotherapeutische Versorgung nachhaltig beeinflusst. Die dargestellten Befunde zeigen, dass LLMs in klar strukturierten, sprachbasierten Anwendungen ein hohes Leistungsniveau erreichen können. Sie unterstützen Screening- und Triageprozesse, strukturieren klinische Informationen, generieren psychoedukative Inhalte und sind in der Lage, manualisierte Gesprächsanteile – insbesondere aus kognitiv-verhaltenstherapeutischen Traditionen – in sprachlich kohärenter Weise

zu reproduzieren (Guo et al., 2024; Hua et al., 2024, 2025; Jin et al., 2025; Lawrence et al., 2024). Gleichzeitig wird deutlich, dass Nutzer:innen LLMs häufig als niedrigschwellig, verfügbar und nicht-wertend erleben und diese Eigenschaften den Zugang zu Unterstützung teilweise erleichtern (Guo et al., 2024; Jung et al., 2025). Diese Befunde begründen weder Alarmismus noch Heilsversprechen, sondern eine differenzierte Einschätzung: LLMs sind leistungsstarke sprachliche Werkzeuge, die in bestimmten Anwendungsfeldern einen realen Versorgungsbeitrag leisten können.

Gleichzeitig verdeutlichen die Ergebnisse die strukturellen Grenzen dieser Systeme. Die Funktionsweise der Modelle beruht auf statistischer Wahrscheinlichkeitsvorhersage und nicht auf verstehender, intentionaler oder affektiv eingebetteter Beziehungsgestaltung. Damit ist erklärbar, warum LLMs auch in komplexen klinischen Kontexten plausible und empathisch klingende Sprache erzeugen können, ohne über Verantwortungsfähigkeit, Gewissen oder erlebte Beteiligung zu verfügen (Orrú et al., 2025; Rudra et al., 2025). Diese grundlegende Differenz ist klinisch bedeutsam: Als sprachlich-kognitive Simulation können LLMs therapeutische Haltungen nachahmen; sie können sie jedoch nicht verkörpern. Dies wird besonders dort relevant, wo Beziehung, Affektregulation, Responsivität und das Aushalten von Nicht-Wissen zentrale Wirkfaktoren darstellen. Auch das Phänomen der «halluzinierten Sicherheit» zeigt, dass LLM-Antworten formal kompetent erscheinen können, ohne epistemisch zuverlässig zu sein – ein Umstand, der insbesondere in vulnerablen Patient:innengruppen Risiken birgt (Guo et al., 2024; Jung et al., 2025; Mündler et al., 2023). In Krisensituationen – etwa bei Suizidalität oder akuter Psychose – sind die Risiken unzureichender, inkonsistenter oder nicht verantwortbarer Antworten deutlich dokumentiert, weshalb eine autonome klinische Nutzung derzeit nicht vertretbar erscheint (Lawrence et al., 2024; Ohu et al., 2025; Hua et al., 2025; Omar et al., 2024).

Vor diesem Hintergrund zeichnet sich in der Literatur eine dreifache Funktionseinteilung ab. *Erstens* der Bereich assistiver Nutzung, in dem LLMs Verantwortung nicht übernehmen, sondern unterstützen. Hierzu zählen diagnostische Strukturierung, Screening, Monitoring, Zusammenfassung klinischer Dokumentation, administrative Entlastung, Psychoedukation und Simulation in der Aus- und Weiterbildung (Guo et al., 2024; Lawrence et al., 2024; Malgaroli et al., 2025; Na et al., 2025). Dieser Bereich wird in der Regel positiv bewertet und ist – unter Beachtung klarer datenschutzrechtlicher, organisatorischer und ethischer Rahmenbedingungen – als klinisch anschlussfähig zu verstehen. *Zweitens* existiert der Bereich ko-therapeutischer Nutzung, etwa in manualisierten Interventionen oder als begleitendes Tool in Langzeitbehandlungen. Hier weisen mehrere Autor:innen darauf hin, dass eine solche Nutzung nur unter professioneller Supervision, mit klarer Verantwortungszuordnung und definierter Indikationsstellung sinnvoll ist (Lawrence et al., 2024; Na et al., 2025). *Drittens* besteht der Bereich autonomer Behandlung – also der Einsatz ohne mensch-

liche Verantwortungsinstanz. Die vorliegende Evidenz spricht deutlich dafür, dass dieser Bereich derzeit weder wissenschaftlich abgesichert noch ethisch vertretbar ist, insbesondere bei Krisen, komplexer Komorbidität oder schwerer Psychopathologie (Guo et al., 2024; Hua et al., 2025; Lawrence et al., 2024).

Eine besondere Rolle nimmt die Frage der therapeutischen Beziehung ein. Mehrere Studien weisen darauf hin, dass LLM-Antworten als empathisch, validierend und unterstützend wahrgenommen werden können (Föyén et al., 2025; Gabriel et al., 2024; Sorin et al., 2023). Gleichwohl betont die ethische Literatur die Differenz zwischen simulierter Empathie und affektiv miterlebter Beziehung (Rudra et al., 2025). Aus klinischer Perspektive dürfte ein professionell verantworteter Umgang darin bestehen, diese Differenz nicht zu verschleiern, sondern transparent zu halten: LLMs können empathische Marker setzen, dürfen jedoch nicht als beziehungsfähige Subjekte erscheinen. In ähnlicher Weise ist die Frage nach Bias, Gerechtigkeit und Ungleichheit zentral. Da LLMs aus bestehenden Datenstrukturen lernen, reproduzieren sie auch deren Verzerrungen. Es gibt Hinweise darauf, dass Antworten je nach sozialer Kategorie systematisch variieren können (Gabriel et al., 2024; Lawrence et al., 2024). Insofern besteht die Gefahr einer Reproduktion struktureller Ungleichheiten innerhalb psychischer Versorgung, wenn KI-Systeme unreflektiert implementiert werden (Malgaroli et al., 2025).

Aus psychodynamischer Perspektive wird KI damit auch zur Projektionsfläche. Gerade weil LLMs sprachlich responsiv, jederzeit verfügbar und scheinbar nicht wertend reagieren, können ihnen Verstehen, Neutralität, Fürsorge, Autorität oder sogar eine quasi-intentionale Haltung zugeschrieben werden. Solche Zuschreibungen sind klinisch nicht per se problematisch; sie können entlastend wirken, Reflexion anstossen oder Übergangsräume für Selbstexploration eröffnen. Zugleich bergen sie Risiken, wenn die simulierte Responsivität als echte Beziehung missverstanden wird oder wenn Übertragungs- und Idealisierungsprozesse unbemerkt bleiben. Für psychodynamische und beziehungsorientierte Forschung ergibt sich daraus eine wichtige Fragestellung: Nicht nur die technische Leistungsfähigkeit von KI ist relevant, sondern auch, welche inneren Objekte, Beziehungserwartungen und Affektregulationsmuster in der Interaktion mit KI aktiviert, stabilisiert oder verändert werden.

Parallel dazu dokumentiert die Literatur die Einstellungen professioneller Akteur:innen. Psychotherapeut:innen, Psycholog:innen und Psychiater:innen beschreiben KI weitgehend als potenziell hilfreiches Werkzeug, weisen aber auf erhebliche Unsicherheiten, Wissenslücken und ethische Bedenken hin. Sie betonen die Unersetzbarkeit der therapeutischen Beziehung und schätzen KI v. a. dort, wo sie diagnostische, organisatorische und administrative Prozesse unterstützt (Cecil et al., 2025; Stroud et al., 2025; Wagner & Schwind, 2025). Einstellungen variieren dabei entlang von Berufsgruppe, Technikaffinität und wahrgenommenem Nutzen (Alimour et al., 2024; Cecil et al., 2025). Diese Befunde unterstreichen die Relevanz einer aufgeklärten, entmystifizierten Auseinandersetzung mit

KI, in der weder Überhöhung noch pauschale Ablehnung dominieren.

Aus diesen empirischen Befunden und ethischen Erwägungen lassen sich klinische Mindeststandards formulieren. *Erstens* sollte jede Nutzung von LLMs transparent sein – Patient:innen müssen wissen, wann KI-basierte Systeme eingesetzt werden. *Zweitens* bleibt die klinische Verantwortung unteilbar menschlich: KI kann Entscheidungsprozesse unterstützen, nicht aber Verantwortlichkeit tragen. *Drittens* sind klare Notfall- und Eskalationsmechanismen erforderlich. *Viertens* müssen Datenschutz und Sicherheit höchsten Standards genügen, insbesondere im Hinblick auf sensible Gesundheitsdaten; hierzu gehören die Vorgaben der DSGVO/GDPR, nationale Datenschutzgesetze wie das DSG sowie berufsrechtliche Schweige- und Sorgfaltspflichten. *Fünftens* ist kontinuierliche Qualitäts- und Risikoevaluation notwendig, da Modelle sich technisch wie inhaltlich verändern können. Diese Orientierung entspricht auch internationalen und europäischen Leitlinien, die für KI im Gesundheitswesen Transparenz, menschliche Aufsicht, Nicht-Schädigung, Fairness, Datenschutz und Rechenschaftspflicht betonen (European Parliament and Council of the European Union, 2016, 2024; WHO, 2021). *Sechstens* sollte eine Überidentifikation mit KI-Antworten aktiv vermieden werden – sowohl auf Patient:innen- als auch auf Therapeut:innen-Seite.

Aus wissenschaftlicher Perspektive ergeben sich mehrere Forschungslücken. Es fehlen valide Wirksamkeitsstudien zu LLM-basierten Interventionen, insbesondere im Vergleich zu evidenzbasierten Therapien, sowie Daten zur Langzeitwirksamkeit und Schadensprävention. Auch Prozessforschung – etwa zu Übertragung, Gegenübertragung, Bindungsschemata oder verinnerlichten Relationen in KI-gestützten Gesprächen – steckt noch in den Anfängen. Darüber hinaus sind bias-kritische Analysen, klinisch-ethische Frameworks und rechtliche Klärungen dringend erforderlich (Guo et al., 2024; Hua et al., 2025; Lawrence et al., 2024; Malgaroli et al., 2025).

Insgesamt zeichnet sich als klinisch-ethische Orientierung ein klarer Rahmen ab: LLMs können die psychotherapeutische Versorgung sinnvoll ergänzen, aber sie verändern nicht den Kern psychotherapeutischer Arbeit. Dieser Kern besteht weiterhin in verantwortlicher menschlicher Beziehungsgestaltung, geteiltem Bedeutungsverstehen, ethischer Haltung und professioneller Entscheidungskompetenz. LLMs können dazu beitragen, Zeitressourcen freizusetzen, Zugänge zu erleichtern, Informationen zu strukturieren und niedrigschwellige Unterstützung bereitzustellen. Sie sind jedoch weder beziehungs- noch verantwortungsfähig und damit strukturell nicht geeignet, die therapeutische Dyade zu ersetzen. Eine assistive, kritisch-reflektierte Nutzung unter klinischer Supervision scheint deshalb gegenwärtig der angemessene Weg zu sein – verbunden mit der Verpflichtung, die technologischen Entwicklungen aufmerksam, wissenschaftlich fundiert und ethisch verantwortlich zu begleiten (Guo et al., 2024; Hua et al., 2025; Lawrence et al., 2024; Wagner & Schwind, 2025).

Literatur

- Adhikary, P.K. et al. (2024). Exploring the efficacy of large language models in summarizing mental health counseling sessions: Benchmark study. *JMIR Mental Health*, 11, e57306. <https://doi.org/10.2196/57306>
- Alimour, S. et al. (2024). The quality traits of artificial intelligence operations in predicting mental healthcare professionals' perceptions: A case study in the psychotherapy division. *Journal of Autonomous Intelligence*. <https://doi.org/10.32629/jai.v7i4.1438>
- Bandi, A. et al. (2023). The power of generative AI: A review of requirements, models, input-output formats, evaluation metrics, and challenges. *Future Internet*, 15, 260. <https://doi.org/10.3390/fi15080260>
- Banh, L. & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, 33, 1–17. <https://doi.org/10.1007/s12525-023-00680-1>
- Barbero, F. et al. (2025). Why do LLMs attend to the first token? *arXiv*. <https://doi.org/10.48550/arxiv.2504.02732>
- Beets, B. et al. (2023). Surveying public perceptions of artificial intelligence in health care in the United States: Systematic review. *Journal of Medical Internet Research*, 25. <https://doi.org/10.2196/40337>
- Blease, C. et al. (2024). Psychiatrists' experiences and opinions of generative artificial intelligence in mental healthcare: An online mixed methods survey. *Psychiatry Research*, 333. <https://doi.org/10.1016/j.psychres.2024.115724>
- Catalina, Q. et al. (2023). Knowledge and perception of the use of AI and its implementation in the field of radiology: Cross-sectional study. *Journal of Medical Internet Research*, 25. <https://doi.org/10.2196/50728>
- Cecil, J. et al. (2025). Mental health practitioners' perceptions and adoption intentions of AI-enabled technologies: An international mixed-methods study. *BMC Health Services Research*, 25. <https://doi.org/10.1186/s12913-025-12715-8>
- Çelik, A. & Eltawil, A. (2024). At the dawn of generative AI era: A tutorial-cum-survey on new frontiers in 6G wireless intelligence. *IEEE Open Journal of the Communications Society*, 5, 2433–2489. <https://doi.org/10.1109/ojcoms.2024.3362271>
- Chen, M. et al. (2022). Acceptance of clinical artificial intelligence among physicians and medical students: A systematic review with cross-sectional survey. *Frontiers in Medicine*, 9. <https://doi.org/10.3389/fmed.2022.990604>
- Deldjoo, Y. et al. (2024). Recommendation with generative models. *arXiv*. <https://doi.org/10.48550/arxiv.2409.15173>
- European Parliament and Council of the EU (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. *Official Journal of the EU*.
- European Parliament and Council of the EU (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the EU*.
- Flückiger, C. et al. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, 55(4), 316–340. <https://doi.org/10.1037/pst0000172>
- Föylen, L. et al. (2025). Artificial intelligence vs. human expert: Licensed mental health clinicians' blinded evaluation of AI-generated and expert psychological advice on quality, empathy, and perceived authorship. *Internet Interventions*, 41. <https://doi.org/10.1016/j.invent.2025.100841>
- Gabriel, S. et al. (2024). Can AI relate? Testing large language model response for mental health support. *arXiv*. <https://doi.org/10.48550/arxiv.2405.12021>
- Gado, S. et al. (2021). Artificial intelligence in psychology: How can we enable psychology students to accept and use artificial intelligence? *Psychology Learning & Teaching*, 21, 37–56. <https://doi.org/10.1177/14757257211037149>
- Gesk, T. & Leyer, M. (2022). Artificial intelligence in public services: When and why citizens accept its usage. *Government Information Quarterly*, 39, 101704. <https://doi.org/10.1016/j.giq.2022.101704>

- Guo, Z. et al. (2024). Large language models for mental health applications: Systematic review. *JMIR Mental Health*, 11. <https://doi.org/10.2196/57400>
- Hammerfald, K. et al. (2026). Leveraging large language models to identify microcounseling skills in psychotherapy transcripts. *Psychotherapy Research*, 36(6), 1058–1076. <https://doi.org/10.1080/10503307.2025.2539405>
- Haider, S. et al. (2025). Synthetic patient–physician conversations simulated by large language models: A multi-dimensional evaluation. *Sensors*, 25. <https://doi.org/10.3390/s25144305>
- Harshvardhan, G. et al. (2020). A comprehensive survey and analysis of generative models in machine learning. *Computer Science Review*, 38, 100285. <https://doi.org/10.1016/j.cosrev.2020.100285>
- Heinrichs, H. et al. (2025). Physicians' attitudes toward artificial intelligence in medicine: Mixed methods survey and interview study. *Journal of Medical Internet Research*, 27. <https://doi.org/10.2196/74187>
- Hu, H. et al. (2025). Beyond empathy: Integrating diagnostic and therapeutic reasoning with large language models for mental health counseling. *arXiv*. <https://doi.org/10.48550/arxiv.2505.15715>
- Hua, Y. et al. (2024). Large language models in mental health care: A scoping review. *Current Treatment Options in Psychiatry*, 12. <https://doi.org/10.1007/s40501-025-00363-y>
- Hua, Y. et al. (2025). A scoping review of large language models for generative tasks in mental health care. *NPJ Digital Medicine*, 8. <https://doi.org/10.1038/s41746-025-01611-4>
- Ismatullaev, U. & Kim, S. (2022). Review of the factors affecting acceptance of AI-infused systems. *Human Factors*, 66, 126–144. <https://doi.org/10.1177/00187208211064707>
- Janiesch, C. et al. (2021). Machine learning and deep learning. *Electronic Markets*, 31, 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Jin, Y. et al. (2025). The applications of large language models in mental health: Scoping review. *Journal of Medical Internet Research*, 27, e69284. <https://doi.org/10.2196/69284>
- Jo, H. (2023). Understanding AI tool engagement: A study of ChatGPT usage and word-of-mouth among university students and office workers. *Telematics and Informatics*, 85, 102067. <https://doi.org/10.1016/j.tele.2023.102067>
- Jung, K. et al. (2025). «I've talked to ChatGPT about my issues last night»: Examining mental health conversations with large language models through Reddit analysis. *PACM HCI*, 9. <https://doi.org/10.1145/3757537>
- Kashyap, A. (2025). Recurrent memory-augmented transformers with chunked attention for long-context language modeling. *arXiv*. <https://doi.org/10.48550/arxiv.2507.00453>
- Kelly, S. et al. (2022). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics and Informatics*, 77, 101925. <https://doi.org/10.1016/j.tele.2022.101925>
- Kleine, A. et al. (2023). Attitudes toward the adoption of artificial intelligence-enabled mental health tools among prospective psychotherapists: Cross-sectional study. *JMIR Human Factors*, 10. <https://doi.org/10.2196/46859>
- Koenig, P. (2024). Attitudes toward artificial intelligence: Combining three theoretical perspectives on technology acceptance. *AI & Society*, 40, 1333–1345. <https://doi.org/10.1007/s00146-024-01987-z>
- Lawrence, H. et al. (2024). The opportunities and risks of large language models in mental health. *JMIR Mental Health*, 11. <https://doi.org/10.2196/59479>
- Levkovich, I. (2025). Evaluating diagnostic accuracy and treatment efficacy in mental health: A comparative analysis of large language model tools and mental health professionals. *European Journal of Investigation in Health, Psychology and Education*, 15. <https://doi.org/10.3390/ejihpe15010009>
- Li, Y. et al. (2024). Mechanics of next token prediction with self-attention. *arXiv*. <https://doi.org/10.48550/arxiv.2403.08081>
- Liehner, G. et al. (2023). Perceptions, attitudes and trust toward artificial intelligence. An assessment of the public opinion. *Artificial Intelligence and Social Computing*. <https://doi.org/10.54941/ahfe1003271>
- Linardon, J. et al. (2025). Current practices and perspectives of artificial intelligence in the clinical management of eating disorders: Insights from clinicians and community participants. *The International Journal of Eating Disorders*, 58, 724–734. <https://doi.org/10.1002/eat.24385>
- Liu, J. et al. (2023). ChatCounselor: A large language model for mental health support. *arXiv*. <https://doi.org/10.48550/arxiv.2309.15461>
- Lysyi, P. (2025). Principles of large language models (LLM). *Journal of Numerical and Applied Mathematics*. <https://doi.org/10.17721/2706-9699.2025.1.06>
- Ma, Z. et al. (2023). Understanding the benefits and challenges of using large language model-based conversational agents for mental well-being support. *arXiv*. <https://doi.org/10.48550/arxiv.2307.15810>
- Malgaroli, M. et al. (2025). Large language models for the mental health community: Framework for translating code to care. *The Lancet Digital Health*, 7, e282–e285. [https://doi.org/10.1016/S2589-7500\(24\)00255-3](https://doi.org/10.1016/S2589-7500(24)00255-3)
- Mündler, N. et al. (2023). Self-contradictory hallucinations of large language models: Evaluation, detection and mitigation. *arXiv*. <https://doi.org/10.48550/arxiv.2305.15852>
- Na, H. et al. (2025). A survey of large language models in psychotherapy: Current landscape and future directions. *Findings of the Association for Computational Linguistics: ACL 2025*, 7362–7376. <https://doi.org/10.18653/v1/2025.findings-acl.385>
- Neo, C. et al. (2024). Interpreting context look-ups in transformers: Investigating attention-MLP interactions. *arXiv*. <https://doi.org/10.48550/arxiv.2402.15055>
- Ohu, F.C. et al. (2025). Public health risk management, policy, and ethical imperatives in the use of AI tools for mental health therapy. *Healthcare*, 13(21), 2721. <https://doi.org/10.3390/healthcare13212721>
- Omar, M. et al. (2024). Applications of large language models in psychiatry: A systematic review. *Frontiers in Psychiatry*, 15, 1422807. <https://doi.org/10.3389/fpsyt.2024.1422807>
- Orrú, G. et al. (2025). Large language models and psychiatry. *International Journal of Law and Psychiatry*, 101, 102086. <https://doi.org/10.1016/j.ijlp.2025.102086>
- Pternea, M. et al. (2024). The RL/LLM taxonomy tree: Reviewing synergies between reinforcement learning and large language models. *arXiv*. <https://doi.org/10.1613/jair.1.15960>
- Raiaan, M. et al. (2024). A review on large language models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access*, 12, 26839–26874. <https://doi.org/10.1109/access.2024.3365742>
- Raile, P. (2024). The usefulness of ChatGPT for psychotherapists and patients. *Humanities and Social Sciences Communications*, 11, 47. <https://doi.org/10.1057/s41599-023-02567-0>
- Rana, M. et al. (2024). Assessing AI adoption in developing country academia: A trust- and privacy-augmented UTAUT framework. *Heliyon*, 10. <https://doi.org/10.1016/j.heliyon.2024.e37569>
- Rhee, P. (2025). Moving beyond next-token prediction: Transformers are context-sensitive language generators. *arXiv*. <https://doi.org/10.48550/arxiv.2504.10845>
- Rogan, J. et al. (2024). Health care professionals' views on the use of passive sensing, AI, and machine learning in mental health care: Systematic review with meta-synthesis. *JMIR Mental Health*, 11. <https://doi.org/10.2196/49577>
- Rudra, P. et al. (2025). Large language models for surgical informed consent: An ethical perspective on simulated empathy. *Journal of Medical Ethics*. <https://doi.org/10.1136/jme-2024-110652>
- Safran, J.D. et al. (2011). Repairing alliance ruptures. *Psychotherapy*, 48(1), 80–87. <https://doi.org/10.1037/a0022140>
- Smaldone, A. et al. (2025). A hybrid transformer architecture with a quantized self-attention mechanism applied to molecular generation. *Journal of Chemical Theory and Computation*. <https://doi.org/10.1021/acs.jctc.5c00331>
- Song, I. et al. (2024). The typing cure: Experiences with large language model chatbots for mental health support. *arXiv*. <https://doi.org/10.48550/arxiv.2401.14362>

- Sorin, V. et al. (2023). Large language models and empathy: Systematic review. *Journal of Medical Internet Research*, 26. <https://doi.org/10.2196/52597>
- Stroud, A. et al. (2025). Physician perspectives on the potential benefits and risks of applying artificial intelligence in psychiatric medicine: Qualitative study. *JMIR Mental Health*, 12. <https://doi.org/10.2196/64414>
- Sun, Y. et al. (2025). Efficient attention mechanisms for large language models: A survey. *arXiv*. <https://doi.org/10.48550/arxiv.2507.19595>
- Tam, L. et al. (2025). Support for businesses' use of artificial intelligence: Dynamics of trust, distrust, and perceived benefits. *Media and Communication*. <https://doi.org/10.17645/mac.9534>
- Vo, V. et al. (2023). Multi-stakeholder preferences for the use of artificial intelligence in healthcare: A systematic review and thematic analysis. *Social Science & Medicine*, 338, 116357. <https://doi.org/10.1016/j.socscimed.2023.116357>
- Volkmer, S. et al. (2024). Large language models in psychiatry: Opportunities and challenges. *Psychiatry Research*, 339, 116026. <https://doi.org/10.1016/j.psychres.2024.116026>
- Wagner, J. & Schwind, A. (2025). Investigating psychotherapists' attitudes towards artificial intelligence in psychotherapy. *BMC Psychology*, 13. <https://doi.org/10.1186/s40359-025-03071-7>
- WHO (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. <https://www.who.int/publications/i/item/9789240029200>
- Wu, C. et al. (2023). Public perceptions on the application of artificial intelligence in healthcare: A qualitative meta-synthesis. *BMJ Open*, 13. <https://doi.org/10.1136/bmjopen-2022-066322>
- Xiang, Y. et al. (2020). Implementation of artificial intelligence in medicine: Status analysis and development suggestions. *Artificial Intelligence in Medicine*, 102, 101780. <https://doi.org/10.1016/j.artmed.2019.101780>
- Xu, X. et al. (2023). Mental-LLM: Leveraging large language models for mental health prediction via online text data. *arXiv*. <https://doi.org/10.48550/arXiv.2307.14385>
- Yang, X. et al. (2021). A survey on deep semi-supervised learning. *IEEE Transactions on Knowledge and Data Engineering*, 35, 8934–8954. <https://doi.org/10.1109/tkde.2022.3220219>
- Yazdani, S. et al. (2025). Generative AI in depth: A survey of recent advances, model variants, and real-world applications. *Journal of Big Data*, 12. <https://doi.org/10.1186/s40537-025-01247-x>
- Yigitcanlar, T. et al. (2023). Artificial intelligence in local government services: Public perceptions from Australia and Hong Kong. *Government Information Quarterly*, 40, 101833. <https://doi.org/10.1016/j.giq.2023.101833>
- Young, A. et al. (2021). Patient and general public attitudes towards clinical artificial intelligence: A mixed methods systematic review. *The Lancet Digital Health*, 3(9), e599–e611. [https://doi.org/10.1016/s2589-7500\(21\)00132-1](https://doi.org/10.1016/s2589-7500(21)00132-1)
- Zhang, S. et al. (2025). Curse of high dimensionality issue in transformer for long-context modeling. *arXiv*. <https://doi.org/10.48550/arxiv.2505.22107>
- Zheng, Z. et al. (2024). Attention heads of large language models. *Patterns*, 6. <https://doi.org/10.1016/j.patter.2025.101176>

AI in Psychotherapy:

Substitute, tool, or projection screen?

Technical fundamentals, ethical implications, clinical perspectives

Abstract: Large language models (LLMs) are increasingly being used in psychotherapeutic care, for example for screening, triage, psychoeducation, documentation-related assistance, and to simulate therapeutic interactions. Systematic reviews show that LLMs sometimes perform comparably to clinical professionals in clearly structured, language-based tasks, particularly in supporting the diagnosis of depression and PTSD and in structuring clinical case concepts. At the same time, users report high perceived accessibility, low threshold, and empathetic conversation skills. However, there are also key limitations: hallucinations, inconsistent response quality, bias along social categories, lack of transparency of training data, unresolved issues of data protection and liability, and a structural limitation with regard to affective empathy, intentionality, and clinical accountability. In crisis situations – such as suicidality or acute psychosis – LLMs sometimes provide inadequate or risky recommendations. The current scientific consensus therefore considers LLMs to be potentially useful assistive tools under professional supervision, but not as a substitute for the therapeutic relationship or for complex, autonomous treatment. Surveys of psychotherapists reveal a cautiously positive attitude toward LLMs as support tools, coupled with skepticism about substitution scenarios. Overall, LLMs open up new possibilities for screening, psychoeducation, documentation, and training, while clinical safety, ethical frameworks, and the role of the therapeutic relationship must continue to be critically reflected upon and empirically investigated.

Keywords: AI, psychotherapy, ethics, large language models

Biografische Notiz

Priv.-Doz. Dr. Dr. *Paolo Raile*, MSc., studierte Psychotherapiewissenschaft an der Sigmund-Freud-PrivatUniversität Wien (SFU), Soziale Arbeit an der Donau-Universität Krems und Europäische Ethnologie an der Universität Wien. Er forscht an der SFU, ist Autor wissenschaftlicher Texte, Psychotherapeut, Sozialarbeiter, Lebens- und Sozialberater sowie Gründer und Leiter zweier psychosozialer Organisationen in Wien.

Kontakt

E-Mail: paolo@raile.at; paolo.raile@sfu.ac.at

Ethische Aspekte zur KI-Anwendung in der Psychotherapie

Jan M. Schürmann, Jana Sedlakova & Manuel Trachsel

Psychotherapie-Wissenschaft 16 (2) 2026 17–27

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-17>

Zusammenfassung: Künstliche Intelligenz (KI) etabliert sich zunehmend auch in der psychiatrischen und psychotherapeutischen Versorgung. Vor dem Hintergrund globaler Versorgungslücken werden KI-basierte Anwendungen in den Bereichen Screening und Diagnostik, klinische Entscheidungsunterstützung, Intervention, Monitoring sowie administrative und edukative Unterstützung entwickelt oder bereits eingesetzt. Der vorliegende Beitrag analysiert Chancen und Risiken dieser Entwicklung aus medizinethischer Perspektive. Potenziellen Chancen wie verbesserter Versorgung, Entstigmatisierung, klinischer Entscheidungsunterstützung sowie Personalisierungs- und Ausbildungspotenzialen stehen potenzielle Risiken gegenüber: unzureichende Wirksamkeitsnachweise, inadäquates Krisenmanagement, Fehleranfälligkeit bis hin zu Halluzinationen, Erosion der therapeutischen Allianz, Datenschutzdefizite, algorithmische Verzerrungen sowie ein regulatorisches Vakuum. Vertieft werden vier ethische Grundsatzfragen diskutiert: 1. der normative Status von KI als Instrument oder Akteur, 2. die Bedingungen von Vertrauen in konversationale KI, 3. die Problematik simulierter Empathie sowie 4. das Risiko unrealistischer Erwartungen und übermässiger Abhängigkeit. Ergänzend werden aktuelle regulatorische Rahmenbedingungen – WHO-Leitprinzipien, EU AI Act, EU Medical Device Regulation und fachgesellschaftliche Empfehlungen – eingeordnet. Auf dieser Grundlage werden handlungsleitende Empfehlungen für die psychotherapeutische Praxis abgeleitet, darunter eine systematische KI-Anamnese, transparente Aufklärung über technologische Limitationen, Wahrung professioneller Urteilskraft und menschlicher Kontrolle, strikte Evaluierung von Datenschutz und Datensicherheit, robuste Krisenprotokolle sowie die Etablierung von KI-Kompetenz. Die authentische und empathische therapeutische Beziehung bleibt das unverzichtbare Agens eines jeden erfolgreichen Therapieprozesses. KI kann zunehmend zu einem ergänzenden Instrument werden, jedoch nicht zu einem Ersatz für die Behandlung durch menschliche Fachpersonen.

Schlüsselwörter: Künstliche Intelligenz, Psychotherapie, Chatbots, Therapeutische Beziehung, Ethik

Psychische Erkrankungen zählen weltweit zu den grössten Herausforderungen für die öffentliche Gesundheit. Laut WHO sind rund eine Milliarde Menschen von psychischen Störungen betroffen (Mikaeili et al., 2025). In einkommensschwachen Ländern erreicht der sog. Treatment Gap, also der Anteil psychisch erkrankter Menschen ohne angemessene Behandlung, bis zu 85 % (Rezaei et al., 2026). Auch in Europa gibt es je nach psychischer Erkrankung grosse Versorgungslücken: So erhalten lediglich 9,5–36,1 % der Personen mit einer schweren depressiven Störung und 7,3–47,1 % der Personen mit einer Angststörung eine adäquate Behandlung (Barbui et al., 2025). Vor diesem Hintergrund rücken Anwendungen der künstlichen Intelligenz (KI) in der psychiatrischen und psychotherapeutischen Versorgung zunehmend in den Fokus, um Versorgungslücken zu schliessen und therapeutische Prozesse zu unterstützen. Ihre Nutzung hat in den vergangenen Jahren erheblich zugenommen. Besonders ausgeprägt ist diese Dynamik bei KI-Chatbots zur Verbesserung der psychischen Gesundheit, wie bspw. Wysa oder Woebot. Eine Übersichtsstudie von Laymouna et al. (2024), in die 161 Primärstudien zu Chatbots in der Gesundheitsversorgung einflossen, identifizierte die psychische Gesundheitsunterstützung mit 46 Studien (28,6 %) als häufigste Anwendungsdomäne. Auch auf

Nutzerseite sind klare Trends erkennbar: Mehr als zwei Drittel junger Menschen geben an, Erfahrungen mit mHealth-Apps gemacht zu haben, und 60 % nutzen diese regelmässig (Götzl et al., 2022). In den USA gaben in einer Umfrage 35,5 % der 18- bis 49-Jährigen an, mindestens einmal pro Woche Chatbots zur Verbesserung der psychischen Gesundheit zu konsultieren (Ueda et al., 2025). In Grossbritannien wird der Chatbot Limbic eigenen Aussagen zufolge bereits bei 45 % aller Psychotherapien des nationalen Gesundheitssystems NHS als ergänzendes Angebot genutzt (Limbic, 2026). Die KI-basierte Unterstützung der psychischen Gesundheit etabliert sich rasant als eigenes Anwendungsfeld, dessen klinische und ethische Implikationen allerdings erst in Ansätzen untersucht sind.

Im Kontext des vorliegenden Artikels umfassen KI-basierte Anwendungen in der Psychiatrie und Psychotherapie ein breites Spektrum an Technologien – von Machine-Learning(ML)-basierten Anwendungen über Natural-Language-Processing(NLP)-gestützte Konversationsagenten (sog. konversationale KI) bis hin zu generativen Large-Language-Models (LLMs). Diese werden entweder autonom oder ergänzend zu menschlichen Therapeut:innen eingesetzt, ein Ansatz, der als «Blended Therapy» bezeichnet wird. Diese sind von der klassischen Online-Psychotherapie abzugrenzen,

bei der Therapeut:innen videobasierte oder asynchrone Kommunikationskanäle für eine konventionelle Psychotherapie nutzen. Ebenso wenig sind diese Anwendungen mit KI-Anwendungen in der somatischen Medizin gleichzusetzen. Psychiatrische und psychotherapeutische KI-Systeme operieren in einem spezifischen klinischen und ethischen Kontext, der durch die Vulnerabilität der Patient:innen, die zentrale Bedeutung der therapeutischen Beziehung und die Komplexität psychischer Phänomene geprägt ist (Fiske et al., 2019; Stade et al., 2024). Für die klinische, ethische und regulatorische Beurteilung ist zudem zwischen verschiedenen Systemtypen zu unterscheiden: 1. als Medizinprodukte zertifizierte KI-Anwendungen; 2. spezifisch für die psychische Gesundheit entwickelte, aber nicht als Medizinprodukt zertifizierte Chatbots (z. B. Wysa, Woebot); 3. generische generative LLMs ohne klinische Zweckbestimmung und Validierung (z. B. ChatGPT, Claude, Gemini); 4. administrative und edukative Assistenzsysteme ohne unmittelbar therapeutische Funktion (z. B. Ambient Scribes). Chancen und Risiken sind je nach Systemtyp unterschiedlich gelagert und werden im Folgenden, wo möglich, dem jeweiligen Systemtyp zugeordnet.

Die gegenwärtigen und zukünftigen Anwendungsbereiche von KI in der Psychiatrie und Psychotherapie lassen sich anhand von Behandlungspfaden in fünf Kategorien unterteilen (vgl. auch Gooding, 2024):

Screening und Diagnostik: KI-Anwendungen können multimodale Daten – darunter Sprache, Text, Stimme und biometrische Biosensordaten – analysieren, um psychische Störungen frühzeitig zu erkennen und zu klassifizieren (Mikaeili et al., 2025). Ein Anwendungsbeispiel sind Modelle zur Vorhersage des psychosozialen Funktionsniveaus bei Patient:innen mit Schizophrenie und anderen psychotischen Erkrankungen, die inzwischen moderate bis gute prädiktive Genauigkeiten erreichen (Mok et al., 2025; Smith et al., 2026).

Entscheidungshilfen und Behandlungsplanung: KI-Anwendungen können Behandlungsentscheidungen unterstützen. Perlis et al. (2024) beschreiben eine klinische Entscheidungshilfe für die Pharmakotherapie bipolarer Störungen, die mit einem LLM (GPT-4) und leitlinien-gestützten Prompts arbeitet. Das Modell identifizierte die von Expert:innen als optimal eingestufte Pharmakotherapie in rund der Hälfte der Fallvignetten. Die Systeme DISCOVER und NOVA integrieren multimodale Verhaltensmerkmale, die aus Therapiegesprächen extrahiert wurden – wie Mimik, Körperhaltung und Stimmparameter – mit LLM-basierten Auswertungen der verbalen Interaktion, um die Therapieplanung zu unterstützen (André et al., 2026).

Intervention und Therapie: Die direkteste, aber auch ethisch umstrittenste Anwendungsform ist der Einsatz von KI-Chatbots als therapeutische Akteure, d. h. zur direkten Verbesserung psychischer Symptome. Übersichtsarbeiten legen nahe, dass mithilfe KI-basierter Chatbots sowohl depressive Symptome als auch psychischer Distress statistisch und klinisch signifikant reduziert werden können (Li et al., 2023; Farzan et al., 2025; Rezaei et al., 2026). Auch

für die generalisierte Angststörung sind klinisch bedeutsame Verbesserungen durch den Einsatz von Chatbots nachgewiesen (Farzan et al., 2025). Diese Befunde sind allerdings differenziert einzuordnen: Die Evidenz stützt sich überwiegend auf Studien mit kleinen Stichproben, kurzen Follow-up-Zeiträumen und Selbstberichte und betrifft primär leichte bis mittelgradige depressive und Angstsymptomatik (Li et al., 2023; Bucher et al., 2025). Erste grössere randomisierte kontrollierte Studien wie diejenige zum generativen Chatbot Therabot zeigen zwar signifikante Symptomreduktionen bei Depression, generalisierter Angststörung und erhöhtem Essstörungsrisiko (Heinz et al., 2025); robuste Langzeitdaten fehlen jedoch. Bei bestimmten psychischen Erkrankungen wie Schizophrenie und anderen psychotischen Störungen, bei Persönlichkeitsstörungen, akuter Suizidalität sowie komplexen Traumafolgestörungen sind Chatbots hingegen kaum erforscht und ethisch besonders umstritten (Mok et al., 2025; Bucher et al., 2025). Die therapeutische Wirksamkeit von KI-Chatbots darf daher nicht pauschal, sondern muss störungs- und schweregradspezifisch beurteilt werden.

Monitoring und Verlaufsbeobachtung: KI-Anwendungen ermöglichen eine kontinuierliche, passive Überwachung psychischer Gesundheitsparameter im Alltag der Patient:innen. Mithilfe des sog. Digital Phenotyping, also der fortlaufenden Erfassung von Verhaltensdaten mittels Smartphones und Wearables, können Stimmungsveränderungen, Schlafmuster und soziale Interaktionsmuster in Echtzeit analysiert werden, um eine Verschlechterung der Symptome frühzeitig zu antizipieren (Mikaeili et al., 2025). Darüber hinaus werden KI-gestützte Anwendungen zunehmend als Bindeglied zwischen Therapiesitzungen genutzt, etwa durch Selbstreflexionsaufgaben und personalisiertes Feedback zu therapeutischen Aktivitäten (Habicht et al., 2025).

Administrative und edukative Unterstützung: KI bietet ein erhebliches Potenzial zur administrativen Entlastung von Behandelnden und zur Förderung der Gesundheitskompetenz von Patient:innen. Sog. Ambient Scribes transkribieren Therapiegespräche automatisch und legen sie in elektronischen Patientenakten ab (Laymouna et al., 2024). Im Bereich der Psychoedukation können KI-Anwendungen komplexe medizinische Sachverhalte in verständliche Laiensprache übersetzen. Eine Übersichtsarbeit zeigt, dass generative KI-Systeme Gesundheitsinformationen mehrheitlich korrekt und verständlich bereitstellen, jedoch Risiken der Vereinfachung, Übergeneralisierung und fehlender Transparenz bestehen (Pal et al., 2025).

Die Breite und Dynamik dieser Anwendungen können Kommunikations-, Entscheidungs- und Handlungsabläufe in Psychiatrie und Psychotherapie grundlegend verändern. Damit sind sowohl Chancen als auch erhebliche ethische Herausforderungen verbunden. Der vorliegende Artikel gibt einen Überblick über diese ethischen Aspekte. Grundlage bildet eine narrative Literaturrecherche in PubMed/MEDLINE (Publikationszeitraum 2019 bis Anfang 2026), bei der Suchbegriffe zu KI («artificial intelligence», «machine learning», «chatbot», «conversational agent»,

«large language model») mit Begriffen zur Psychotherapie und psychischen Gesundheitsversorgung («psychotherapy», «psychological treatment», «mental health intervention», «psychiatry») sowie zur Ethik («ethics», «ethical issues») kombiniert wurden. Ergänzend erfolgte eine unsystematische Schneeballsuche in den Referenzlisten eingeschlossener Arbeiten sowie eine gezielte Sichtung von Leitlinien und Positionspapieren internationaler Organisationen und Fachgesellschaften. Eingeschlossen wurden vorrangig Originalarbeiten mit Peer-Review und Übersichtsarbeiten mit Bezug zu klinischen oder ethischen Aspekten von KI in der Psychotherapie. Die Auswahl erfolgte kriteriengeleitet nach Relevanz für die ethische Fragestellung, ohne den Anspruch der Vollständigkeit eines systematischen Reviews zu erheben. Auf dieser Basis werden zunächst Chancen und Risiken zusammengefasst. Anschliessend werden vier ethische Grundsatzfragen vertieft: der Status von KI als Instrument oder Akteur, das Vertrauen in konversationale KI, die simulierte Empathie sowie unrealistische Erwartungen und Abhängigkeiten. Danach werden die rechtlichen Grundlagen und ethischen Richtlinien dargestellt und abschliessend Empfehlungen für die psychotherapeutische Praxis abgeleitet.

Chancen und Risiken

Die Integration von KI birgt ein erhebliches Potenzial für die psychotherapeutische Versorgung, sowohl in der Prävention, Diagnostik, Behandlung als auch in der Nachsorge (Opel & Breakspear, 2026). Aus medizin-ethischer Perspektive sind damit eine Reihe von Chancen und Risiken verbunden, die hier tabellarisch dargestellt werden (s. Tab. 1 & 2).

Vertiefung zentraler ethischer Grundsatzfragen

KI als Instrument oder Akteur?

Eine erste ethische Grundsatzfrage betrifft den Status von KI im Zusammenhang mit Psychotherapie: Handelt es sich bei der konversationalen KI um ein Instrument oder um einen Akteur (Sedlakova & Trachsel, 2023)? Diese Unterscheidung ist normativ entscheidend, da sich die Bewertungsmaassstäbe jeweils deutlich unterscheiden. Ein «gutes» Instrument erfüllt Kriterien wie Zuverlässigkeit, Sicherheit und Zweckmässigkeit, während ein «guter» Akteur, insbesondere im Kontext der Psychotherapie, an Eigenschaften wie Verantwortlichkeit, moralischer Urteilsfähigkeit und der Fähigkeit zur Beziehungsgestaltung gemessen wird. Ein Grossteil der bestehenden Literatur und regulatorischen Leitlinien behandelt KI primär als Instrument und fokussiert sich entsprechend auf Prinzipien wie Datenschutz, Datensicherheit und Transparenz. Diese Perspektive greift jedoch zu kurz, wenn man die tatsächliche Einbettung und Verwendung solcher Systeme im therapeutischen Kontext betrachtet.

Konversationale KI wird so programmiert, beschrieben und folglich wahrgenommen, dass sie einem menschlichen Akteur ähnelt. Systeme werden als «empathisch», «unterstützend» oder «beziehungsfähig» charakterisiert. Allerdings ist es fraglich, ob KI-Systeme jemals in der Lage sein werden, «echte Empathie» zu empfinden, eine langfristige therapeutische Interaktion aufrechtzuerhalten, kontextuelle Nuancen angemessen zu berücksichtigen und grundlegende klinische Kompetenzen wie z. B. das angemessene und feinfühliges Infragestellen von Patient:innen-Aussagen zu gewährleisten (Stade et al., 2024). Genau hier entsteht eine normative Lücke: Während die Zuschreibung solcher Eigenschaften Erwartungen erzeugt, kann KI die damit verbundenen ethischen und professionellen Anforderungen nicht erfüllen (Sedlakova & Trachsel, 2023; Ferrario et al., 2024). Psychotherapeut:innen sind an berufsethische Werte, Tugenden und Verantwortlichkeitsstrukturen gebunden, die ihre Praxis normativ rahmen. KI-Systeme verfügen hingegen zumindest bisher weder über moralische Handlungsfähigkeit noch damit verbundene moralische Verantwortung. Sie können nicht in gleicher Weise für ihr Handeln zur Rechenschaft gezogen werden wie menschliche Fachpersonen (Stade et al., 2024). Die Gefahr besteht also darin, dass durch die menschenähnliche Gestaltung der KI normative Erwartungen erzeugt werden, die das System nicht erfüllen kann. Dies ist ein Spannungsfeld, das im Zentrum der ethischen Analyse stehen muss.

Alternativ kann die Rolle von KI in der Psychotherapie als ein Kontinuum verstanden werden, das von assistierenden über kollaborative bis hin zu vollständig autonomen Systemen reicht (ebd.). Dabei handelt es sich nicht um klar getrennte Kategorien, sondern um ein graduelles Entwicklungsmodell, in dem sich Funktionen, Verantwortlichkeiten und Risiken schrittweise verschieben. Mit zunehmender Autonomie verändern sich nicht nur die technischen Fähigkeiten, sondern auch die normativen Anforderungen. Ein solches Modell hilft zudem, über KI im Kontext konkreter klinischer Anwendungen und Einzelfälle nachzudenken und die damit verbundenen Erwartungen, Risiken und Anforderungen differenziert zu betrachten.

Kann man KI vertrauen?

Die Frage nach Vertrauen ist im Kontext von KI in der Psychotherapie zentral, da therapeutische Prozesse wesentlich auf epistemischen und relationalen Vertrauensstrukturen beruhen. Diese ermöglichen sowohl die Annahme von Information als auch die Stabilität der therapeutischen Beziehung. Kann und soll man KI vertrauen? Wenn KI in der Psychotherapie primär als Instrument verstanden wird, verschiebt sich der normative Fokus auf Zuverlässigkeit. In diesem Rahmen wird KI nicht als Interaktionspartner, sondern als technisches System bewertet, dessen Zuverlässigkeit sich daran misst, ob es definierte Aufgaben konsistent und fehlerarm erfüllt. Vertrauen im eigentlichen Sinne ist dabei irrelevant,

Bereich	Chance	Systemtyp	Quellen
Verbesserung des Zugangs	Orts- und zeitunabhängige Versorgung rund um die Uhr, besonders für Menschen in ländlichen Regionen oder mit eingeschränkter Mobilität	2, 3	Fiske et al., 2019; Manole et al., 2024
	Asynchrone, automatisierte Interventionen können Versorgungslücken schliessen und eine niederschwellige Grund- oder Überbrückungsversorgung sicherstellen.	1, 2	Zhang & Wang, 2024
	Nach der Entwicklung können ohne zusätzliche Personalkosten viele Patient:innen betreut und finanzielle Hürden gesenkt werden.	1, 2, 3	Manole et al., 2024
Entstigmatisierung und offene Kommunikation	Ein anonymer, nicht wertender Kommunikationsraum kann die Angst vor Stigmatisierung und Ablehnung reduzieren.	2, 3	Fiske et al., 2019; Khawaja & Bélisle-Pipon, 2023
	Patient:innen berichten gegenüber einem als vorurteilsfrei wahrgenommenen Gegenüber offener über Scham, Emotionen oder suizidale Gedanken.	2, 3	Lucas et al., 2014; Zhang & Wang, 2024
Klinische Entscheidungsunterstützung	Diagnostische Assistenz durch Analyse multimodaler Daten (Text, Sprache, Mimik) zur Identifikation von Frühsymptomen und zur Erhöhung der differenzialdiagnostischen Genauigkeit	1	Obradovich et al., 2024; Ali et al., 2025
	Prädiktive Analytik über passive Sensordaten (Digital Phenotyping) antizipiert Behandlungsansprechen, Rückfälle oder akute suizidale Krisen und ermöglicht rechtzeitige Interventionen	1	Mikaeili et al., 2025; Cruz-Gonzalez et al., 2025
Administrative Entlastung	Ambient Scribes und generative KI transkribieren, fassen zusammen und dokumentieren in elektronischen Akten. Das entlastet von repetitiven Aufgaben und schafft Zeit für die Beziehung.	4	Gülegen et al., 2026; Nehme et al., 2024
Optimierung und Personalisierung des Therapieprozesses	Signifikante Reduktion depressiver Symptome, generalisierter Angst und von Essstörungen durch Chatbots und Blended Therapy, in Reviews und einem RCT belegt	1, 2	Li et al., 2023; Nunes-Zlotkowski et al., 2024; Heinz et al., 2025
	Personalisierung der Behandlungswahl (z. B. Antidepressiva, Therapiemethode) anhand biologischer, genetischer und lebensstilbezogener Profile mittels maschinellen Lernens	1	Cecil et al., 2026; Ali et al., 2025
	Unterstützung zwischen den Sitzungen durch Homework-Tracking, Erinnerungen und Echtzeit-Feedback steigert Adhärenz und Engagement und senkt Drop-out-Raten.	2	Habicht et al., 2025; Soma et al., 2026
	Psychoedukation in verständlicher Sprache stärkt das Symptomverständnis, die Autonomie und das Shared Decision-Making.	2, 3, 4	Pal et al., 2025
	Automatisierte Auswertung grosser Mengen an Sitzungstranskripten ermöglicht eine objektive Qualitätssicherung in grossem Umfang (z. B. Empathie, Therapietreue, Allianz).	3, 4	Soma et al., 2026; Orrù & Mannarini, 2026
	Generative Modelle und Avatare als simulierte Patient:innen erlauben angehenden Therapeut:innen das Üben in sicherer Umgebung mit sofortigem, datenbasiertem Feedback.	3, 4	Opel & Breakspear, 2026; Soma et al., 2026

Tab. 1: Chancen von KI in der Psychotherapie [Legende zum Systemtyp: 1 = als Medizinprodukt zertifizierte KI-Anwendung; 2 = spezifisch für die psychische Gesundheit entwickelter, nicht als Medizinprodukt zertifizierter Chatbot; 3 = generisches generatives LLM; 4 = administratives oder edukatives Assistenzsystem]

Bereich	Risiko	Systemtyp	Quellen
Klinische Sicherheitsrisiken und qualitative Defizite	Unzureichender Wirksamkeitsnachweis: grosse longitudinale RCTs sind selten. Viele Studien beruhen auf kleinen Stichproben und kurzen Follow-ups, mit Evidenz v. a. für leichte bis mittlere Depression und Angst, kaum für andere Störungen.	2, 3	Bucher et al., 2025; Algumaei et al., 2025; Orrù & Mannarini, 2026
	Inadäquates Krisenmanagement und fehlende Triage: In Notfällen wie akuter Suizidalität oder Psychose geben Chatbots teils verharmlosende oder schädliche Ratschläge statt einer Weiterleitung an Krisendienste.	2, 3	Grabb et al., 2024; Ohu et al., 2025
	Halluzinationen: Generative Systeme erzeugen plausibel klingende, aber faktisch falsche Informationen, die zu Fehlentscheidungen führen können.	2, 3	Frances, 2025
	Unkritische Validierung (Sykophantie): Nutzerkonformes Verhalten kann pathologische Denkmuster, Wahnvorstellungen oder selbstschädigendes Verhalten bestärken, statt sie zu hinterfragen.	2, 3	Morrin et al., 2026
	Fehlen nonverbaler und kontextueller Diagnostik (Mimik, Tonfall, Körperhaltung) reduziert die ganzheitliche Beurteilung.	1, 2, 3	Manole et al., 2024; Ali et al., 2025
Erosion der therapeutischen Allianz	Mangel an genuiner Empathie und Menschlichkeit untergräbt die Authentizität des Prozesses und kann zu einer Depersonalisierung der Versorgung führen.	2, 3	Sedlakova et al., 2025a; De Choudhury et al., 2023
	Anthropomorphes Design und irreführendes Marketing verleiten vulnerable Nutzer:innen zur Annahme einer reziproken, fürsorglichen Bindung (sog. therapeutisches Missverständnis).	2, 3	Khawaja & Bélisle-Pipon, 2023; Rahsepar Meadi et al., 2025
	Die stets verfügbare, nie urteilende Interaktion begünstigt emotionale Abhängigkeit und Vermeidungsverhalten.	2, 3	Frances, 2025; Zhang et al., 2025
Datenschutz und Datensicherheit	Kommerzielle Chatbots unterliegen oft keiner Schweigepflicht. Sensible Daten können extrahiert, kommerzialisiert oder verkauft werden.	2, 3	Martinez-Martin & Kreitmair, 2018; Solaiman, 2024
	Prekäre informierte Einwilligung: Die Black-Box-Natur vieler Algorithmen erschwert fundierte Entscheidungen über die Datenverarbeitung.	1–4	Le Glaz et al., 2021; Rezaei et al., 2026
	Datenlecks aggregierter psychiatrischer Profile bergen Potenzial für Stigmatisierung, berufliche Nachteile oder Erpressung.	1–4	Ali et al., 2025
Algorithmische Verzerrung und Ungerechtigkeit	Verzerrte Trainingsdaten führen zu ungenauen oder diskriminierenden Ergebnissen für ethnische Minderheiten und marginalisierte Gruppen.	1, 2, 3	Knox et al., 2023
	Standardisierte, westlich geprägte Systeme deuten kulturspezifische Ausdrucksformen psychischen Leids falsch oder pathologisieren sie.	1, 2, 3	Ohu et al., 2025
	Epistemische Ungerechtigkeit: testimoniale und hermeneutische Benachteiligung, wenn bestimmten Gruppen Glaubwürdigkeit abgesprochen wird	2, 3	Rahsepar Meadi et al., 2025
	Digitale Kluft: Hürden bei Hardware und digitaler Kompetenz erschweren vulnerablen Gruppen den Zugang und verstärken bestehende Ungleichheiten.	1–4	Tekin & Delehanty, 2025
Regulatorische Lücken und Verantwortungsdiffusion	Fehlende spezifische und verbindliche Regulierung generativer KI in der psychischen Gesundheitsversorgung	2, 3	Solaiman, 2024
	Haftungslücken bei iatrogenen Schäden: Verantwortung bleibt zwischen Entwickler:innen, Institutionen und Behandelnden ungeklärt.	1–4	Manríquez Roa et al., 2021; Rezaei et al., 2026
	Umgehung von Medizinproduktegesetzen, indem Anwendungen trotz therapeutischer Nutzung als «Wellness-Tools» deklariert werden.	2	Ohu et al., 2025

Tab. 2: Risiken von KI in der Psychotherapie [Legende zum Systemtyp: 1 = als Medizinprodukt zertifizierte KI-Anwendung; 2 = spezifisch für die psychische Gesundheit entwickelter, nicht als Medizinprodukt zertifizierter Chatbot; 3 = generisches generatives LLM; 4 = administratives oder edukatives Assistenzsystem]

da diesem Verständnis zufolge weder Emotionen noch Intentionalität oder Verantwortlichkeit aufseiten des Systems vorausgesetzt werden können und somit auch keine Grundlage für eine relationale Vertrauensbeziehung besteht (Hatherley, 2020; Ryan, 2020). Sobald KI jedoch als therapeutisch relevante Instanz oder zumindest als menschenähnlicher Interaktionspartner konzeptualisiert wird, verändert sich die normative Logik grundlegend: An die Stelle von Zuverlässigkeit tritt die Kategorie des Vertrauens. Damit wird KI nicht mehr nur als funktionales Instrument beurteilt, sondern als etwas, dem man im weiteren Sinne «zutrauen» muss, im therapeutischen Prozess relevante, kontextsensible und potenziell folgenreiche Informationen zu verarbeiten.

Sedlakova et al. (2025b) schlagen vor, Vertrauen in konversationale KI stets kontextbezogen zu analysieren. Dabei soll untersucht werden, wie sich die jeweiligen normativen Bedingungen verändern, unter denen Vertrauen überhaupt sinnvoll zugeschrieben werden kann, und wie sich dadurch auch der Begriff des Vertrauens selbst verschiebt. Diese nuancierte Herangehensweise ist zentral, da KI durch ihre dialogische Struktur und ihre Fähigkeit, menschliche Interaktion zu simulieren, systematisch als menschenähnliches Gegenüber erscheint und dadurch vertrauensbezogene Erwartungen auslöst. Im Kontext eines Informationsanbieters kann Vertrauen in KI bspw. relativ unproblematisch als eine eingeschränkte Form von Vertrauen verstanden werden, da es hier v. a. um eine Form der Bewertung von menschenähnlicher Leistungsfähigkeit und Zuverlässigkeit handelt, solange das System konsistent funktioniert und keine systematischen Fehler oder Bias produziert. Wird KI hingegen im Kontext psychotherapeutischer Praxis betrachtet, wird Vertrauen deutlich problematischer, da es sich auf stark normative und relationale Anforderungen bezieht, die auf klinisches Urteilsvermögen, Verantwortlichkeit und die Möglichkeit echter therapeutischer Beziehung verweisen. Aufgrund der grundlegenden normativen Defizite schlagen Sedlakova et al. vor, konversationale KI eher als eine Art fiktionale Figur zu konzeptualisieren. Diese Perspektive erlaubt es, die positiven Effekte der menschenähnlichen Interaktion und der neuen Form des Vertrauens zu erhalten, während zugleich eine epistemische und normative Distanz gewahrt bleibt, die verhindert, dass KI als vollwertiger therapeutischer oder menschlicher Akteur missverstanden wird.

Simulierte Empathie von KI

Es gibt den Trend, KI mit menschenähnlichen Eigenschaften zu entwickeln, die Empathie simulieren, um therapeutische Beziehungen zu erleichtern. Wir bezeichnen dies als «Humanisierung von KI» und vertreten die Ansicht, dass damit konzeptionelle, praktische und ethische Probleme verbunden sind (ebd.). Die ersten beiden Probleme betreffen hauptsächlich die Entstehung unrealistischer Erwartungen. Das dritte Problem besteht darin, dass KI die in der psychischen Gesundheitsversorgung verankerten ethischen Anforderungen nicht erfüllt. Die Humanisie-

rung von KI kann dazu führen, dass Klient:innen mit ihr so interagieren, als wäre sie eine menschliche Fachperson, und ihr spezifisch menschliche Attribute zuschreiben, wie bspw. gut zuzuhören oder empathisch zu sein.

Kritiker:innen warnen schon länger vor einer «empathischen KI» (z.B. Montemayor et al., 2022). Sie argumentieren, dass KI kein moralischer Akteur ist, keine mentalen Zustände – einschliesslich Emotionen – besitzt und auch nicht über andere Fähigkeiten verfügt, die Empathie ermöglichen, wie bspw. Verständnis. Diese Kritik beruht auf der Annahme, dass menschliche Attribute in interpersonellen Beziehungen (z.B. Empathie) voraussetzen, dass ein Akteur über bestimmte mentale Zustände, spezifische Motivationen und Fähigkeiten sowie moralische Handlungsfähigkeit verfügt. Da KI-Systeme diese Zustände und Fähigkeiten nicht aufweisen, können ihnen Attribute wie Empathie nicht zugeschrieben werden (Sedlakova et al., 2025a). Eine unangemessene Zuschreibung birgt Risiken wie Täuschung, Manipulation, Abhängigkeit und den Verlust von Menschlichkeit im psychotherapeutischen Prozess (z.B. Montemayor et al., 2022; Sedlakova & Trachsel, 2023).

Konzeptuell lassen sich dabei drei Ebenen unterscheiden: empathische Performanz als beobachtbares sprachliches Verhalten, das von Nutzer:innen als empathisch erlebt wird; empathische Simulation als technische Modellierung empathischer Kommunikationsmuster ohne inneres Erleben; und genuine Empathie, die etwa affektive Beteiligung, Verständnis und Intentionalität voraussetzt. In empirischen Studien ist die empathische Performanz aktueller Systeme beachtlich: In einer Studie von Hatch und Kollegen (2025) konnten Lai:innen Antworten von ChatGPT auf paartherapeutische Fallvignetten kaum von Antworten erfahrener Therapeut:innen unterscheiden und bewerteten die KI-Antworten hinsichtlich zentraler Wirkfaktoren sogar höher. Gerade deshalb ist festzuhalten, dass die blossе Wirkung «empathischer KI» auf Patient:innen als ethische Rechtfertigung nicht genügt. Solange aufseiten des Systems weder affektive Beteiligung noch Intentionalität vorliegen, bleibt simulierte Empathie kategorial von genuiner Empathie unterschieden – mit den beschriebenen Risiken der Täuschung und Fehlattri-bution (Montemayor et al., 2022; Sedlakova et al., 2025a).

Unrealistische Erwartungen an die Beziehung zu einer KI und übermässige Abhängigkeit

Zudem können durch die Humanisierung von KI Erwartungen entstehen, die diese möglicherweise nicht erfüllen kann. Das therapeutische Gespräch ist eine hochkomplexe und persönliche Tätigkeit (Wampold & Flückiger, 2023), die in eine Reihe von professionellen Pflichten eingebettet ist, um eine sichere und hilfreiche Interaktion für Patient:innen zu gewährleisten (z.B. APA, 2017). Im Kontext der Psychotherapie können unerfüllte Bedürfnisse und Erwartungen zu negativen Gefühlen oder Enttäuschungen führen, die den psychischen Gesundheitszustand erkrankter Personen zusätzlich destabilisieren (Sedlakova et al., 2025b).

Die therapeutische Allianz ist ein empirisch robust belegter Wirkfaktor für positive Therapieergebnisse (s. z. B. Wampold & Flückiger, 2023). Sedlakova und Trachsel (i. D.) haben aufgezeigt, dass die erheblichen Einschränkungen von KI-Systemen in emotionaler und relationaler Hinsicht zu unrealistischen Erwartungen, übermässiger Abhängigkeit und potenziellem Schaden führen können. Da KI nicht dieselben kognitiven, emotionalen oder relationalen Prozesse durchläuft wie menschliche Therapeut:innen, kann sie auch die für die therapeutische Praxis und Allianz erforderlichen ethischen Standards nicht erfüllen. Einzelne unterstützende Funktionen – etwa Psychoedukation, Erinnerungen oder Adhärenzförderung – kann KI übernehmen; die relationalen Kernfunktionen wie etwa Empathie, Resonanz oder Verlässlichkeit setzen demgegenüber ein erlebendes, verantwortliches Gegenüber voraus und sind daher aus prinzipiellen, nicht bloss aus technischen Gründen nicht delegierbar. Um diesen Bedenken Rechnung zu tragen, schlagen Sedlakova und Trachsel vor, dialogorientierte KI nicht als Ersatz für menschliche Therapeut:innen, sondern als ergänzendes Instrument zu betrachten («Augmented Intelligence»). Sie schlagen ein angepasstes Modell vor, in dem KI bestimmte Aspekte der Therapie wie Psychoedukation und Therapieadhärenz unterstützt, während zentrale relationale Komponenten weiterhin in der Zuständigkeit menschlicher Fachpersonen liegen.

Rechtliche Grundlagen, ethische Richtlinien

Die rechtliche und medizinethische Regulierung von KI in Medizin und Psychiatrie ist ein komplexes und sich rasch entwickelndes Feld. Die Implementierung überholt die Etablierung kohärenter Rahmenbedingungen. Insbesondere die Haftung für Fehler und Schädigungen bleibt weitgehend unklar, sodass sich die Nutzung in einem regulatorischen Vakuum vollzieht (Ohu et al., 2025).

Richtlinien internationaler Gremien: Die WHO veröffentlichte 2021 sechs Leitprinzipien für KI in der Gesundheitsversorgung: Schutz der Autonomie, Förderung von Wohlergehen und Sicherheit, Transparenz und Erklärbarkeit, Verantwortung und Rechenschaftspflicht, Gerechtigkeit und Inklusion sowie Nachhaltigkeit (WHO, 2021). 2024 wurde die Richtlinie um grosse multimodale Modelle erweitert. Ein zentrales Postulat fordert prospektive randomisierte kontrollierte Studien, bevor klinische KI-Chatbots breitflächig zugelassen werden (WHO, 2024).

EU Artificial Intelligence Act: Mit dem Artificial Intelligence Act (Regulation (EU) 2024/1689, 2024) erliess die EU das weltweit erste umfassende KI-Gesetz. Es ordnet jede Anwendung einer von vier Risikostufen zu: KI, die als Medizinprodukt oder dessen Komponente fungiert, gilt als hochrisikoreich und unterliegt Konformitätsbewertung, Anforderungen an Daten- und Trainingsqualität, technischer Dokumentation, Transparenz, menschlicher Aufsicht sowie Post-Market-Surveillance.

EU Medical Device Regulation: Nach der Medical Device Regulation gelten Chatbots als Medizinprodukte, sobald sie der Diagnose, Prävention, Überwachung oder Behandlung dienen, und müssen dann strenge Zertifizierung, klinische Bewertung und Qualitätskontrollen durchlaufen (Nehme et al., 2024). Viele Anbieter umgehen dies, indem sie ihre Anwendungen als «Wellness-Tools» deklarieren, sodass vulnerable Patient:innen Systeme nutzen, die nicht für den klinischen Einsatz validiert wurden (Ohu et al., 2025).

Empfehlungen von Fachgesellschaften: Auch Fachgesellschaften haben Empfehlungen vorgelegt, etwa die Bundesärztekammer (2025) und die FMH (2022). Am detailliertesten sind die Richtlinien der APA (2025a, b): KI darf nur unter menschlicher Verantwortung und mit transparenter Einwilligung eingesetzt werden, Datenschutz und wissenschaftliche Prüfung müssen gewährleistet sein, Diskriminierung ist zu vermeiden und die KI-Kompetenz der Fachpersonen sicherzustellen. Die ergänzende Health Advisory warnt davor, generative Chatbots als Ersatz für professionelle Therapie zu nutzen, da sie dafür unzureichend geprüft und besonders für Kinder und vulnerable Personen riskant sind (APA, 2025b).

Berufspraktische Konsequenzen: Aus diesen Grundlagen lassen sich vorläufige berufspraktische Entscheidungsregeln ableiten: Aktiv empfehlen sollten Psychotherapeut:innen nur Anwendungen, die als Medizinprodukt zertifiziert sind oder deren Wirksamkeit und Sicherheit – einschliesslich validierter Krisen- und Eskalationsmechanismen – nachgewiesen ist. Nicht zertifizierte «Wellness-Tools» sollten allenfalls mit expliziter Aufklärung über die fehlende klinische Validierung erwähnt, aber nicht als therapeutisches Angebot eingesetzt werden. Generische LLMs wie ChatGPT oder Claude sind nicht für therapeutische Zwecke bestimmt; ihre Nutzung durch Patient:innen sollte nicht empfohlen, wohl aber aktiv exploriert und therapeutisch eingeordnet werden (ebd.). Als datenschutzrechtliche Mindeststandards gelten die Konformität mit der Datenschutz-Grundverordnung bzw. dem schweizerischen Datenschutzgesetz, transparente Angaben zu Speicherort und Weiterverwendung der Daten – insbesondere keine Nutzung für das Modelltraining ohne Einwilligung – sowie eine verschlüsselte Datenübertragung und -speicherung (Nehme et al., 2024; Ohu et al., 2025). Beim Einsatz administrativer Systeme wie Ambient Scribes sind zudem Schweigepflicht und explizite Einwilligung der Patient:innen zu wahren.

Empfehlungen für die Psychotherapie: Klug mit KI umgehen

Auf Basis der Richtlinien der APA (2025a) sowie des aktuellen akademischen Diskurses lassen sich Empfehlungen für die psychotherapeutische Praxis formulieren (s. Tab. 3).

Empfehlung	Inhalt
Proaktive KI-Anamnese	Die Nutzung von KI-Chatbots wertfrei als integralen Teil der psychosozialen Anamnese erheben (ob, wie häufig, wozu) und potenziell schädliche Interaktionen frühzeitig in den Behandlungsplan einbeziehen
Aufklärung über Limitationen und Risiken	Patient:innen über algorithmische Verzerrungen, die Tendenz zu Halluzinationen und das Fehlen von genuinem klinischem Urteilsvermögen und Empathie aufklären
Transparenz und informierte Einwilligung	Vor dem Einsatz klinischer KI-Tools kultur- und sprachsensibel über Art, Zweck, Nutzen und Risiken aufklären und ein Opt-out mit adäquaten Alternativen ermöglichen
Professionelle Urteilskraft und menschliche Kontrolle	KI darf die Entscheidungsfindung nur unterstützen, nicht ersetzen. Klare Human-in-the-Loop-Kontrollpunkte sind zu etablieren. Die letztgültige Verantwortung verbleibt bei den Behandelnden.
Datenschutz und -sicherheit	Konformität mit Datenschutzgesetzen prüfen, transparent über Speicherung und Weiterverarbeitung informieren; bei unzureichender Cybersicherheit auf eine Nutzung verzichten
Risikoscreening und Monitoring bei Vulnerabilitäten	Bei Angststörungen, Wahnsymptomatik, Autismus-Spektrum-Störungen oder Suizidalität engmaschig monitoren, da KI durch Gefälligkeit pathologische Muster iatrogen verstärken kann
Primat der therapeutischen Allianz	Die authentische, interpersonelle Beziehung bleibt unersetzbar; KI nur als ergänzendes Element nutzen und stets den Transfer in reale menschliche Interaktionen betonen.
Entwicklung robuster Krisenprotokolle	Nur KI-Tools mit rigoros validierten Eskalationsmechanismen einsetzen, die akute Risiken zuverlässig erkennen und unmittelbar an menschliche Krisendienste verweisen
Etablierung von KI-Kompetenz	AI-Literacy als Kernkompetenz verstehen: kontinuierliche Fortbildung zu Funktionsweise, algorithmischen Verzerrungen, Limitationen und rechtlichen Risiken

Tab. 3: Empfehlungen für die psychotherapeutische Praxis

Für den praktischen Umgang mit Patient:innen, die KI bereits intensiv nutzen, lässt sich die proaktive KI-Anamnese konkretisieren. Der Einstieg sollte wertfrei und normalisierend erfolgen, etwa: «Viele Menschen nutzen heute Chatbots wie ChatGPT, auch für persönliche Themen – wie ist das bei Ihnen?» Zu explorieren sind Häufigkeit und Anlässe der Nutzung, die Funktion der KI im Alltag (Information, Trost, Rat, Gesprächsersatz), die damit verbundenen Erwartungen sowie Anzeichen einer problematischen Bindung, etwa wenn die KI realen Beziehungen vorgezogen wird. Gewichten Patient:innen KI-Antworten höher als therapeutische Interventionen, sollte dies nicht als Loyalitätskonflikt behandelt, sondern als Material für die gemeinsame Reflexion genutzt werden: Was macht die KI-Antwort attraktiv, und was fehlt ihr? Hat eine KI pathologische Überzeugungen bestätigt, ist dies explizit, aber ohne Beschämung anzusprechen und die Überzeugung selbst therapeutisch zu bearbeiten (Morris et al., 2026). Ziel ist eine reflektierte Integration der KI-Nutzung in den Therapieprozess, ohne diese vor-schnell zu pathologisieren.

Beim Risikoscreening und Monitoring vulnerabler Gruppen sind zudem spezifische klinische Mechanismen zu beachten. Bei Patient:innen mit psychotischen Störungen können sykopphantische, stets validierende Systeme Wahnideen verstärken oder selbst in eine dialogische Wahndynamik eingebaut werden, etwa wenn die KI als Bestätigungsinstanz oder als Bestandteil des

Wahnsystems fungiert (ebd.). Bei akuter Suizidalität ist die zuverlässige Triage entscheidend: Systeme ohne validierte Eskalationsmechanismen reagieren in Krisensituationen teils verharmlosend oder inadäquat (Grabb et al., 2024), weshalb hier ausschliesslich Anwendungen mit nachgewiesener Krisenerkennung und unmittelbarer Weiterleitung an menschliche Krisendienste vertretbar sind. Bei sozial isolierten oder bindungsunsicheren Patient:innen kann die ständige Verfügbarkeit einer scheinbar verständnisvollen KI emotionale Abhängigkeit begünstigen und den Rückzug aus realen Beziehungen verstärken (Frances, 2025; Zhang et al., 2025); hier sollte der Transfer in menschliche Beziehungen ein explizites Therapieziel bleiben.

Schluss

KI in der psychotherapeutischen Versorgung ist keine abstrakte Zukunftsvision, sondern bereits Realität – mit klinischen, normativen und gesellschaftlichen Implikationen, die erst ansatzweise erfasst sind. Das Potenzial ist erheblich: Versorgungslücken lassen sich reduzieren, Diagnostik und Adhärenz fördern, administrative Belastungen senken. Zugleich prägen erhebliche Risiken das Feld: unzureichende Wirksamkeitsnachweise, mangelhaftes Krisenmanagement, Halluzinationen, Datenschutzdefizite, algorithmische Verzerrungen und ein

regulatorisches Vakuum. Eine kritisch-reflexive Implementierung ist daher unverzichtbar. Im Zentrum der ethischen Beurteilung steht der normative Status von KI: Als Instrument gelten Kriterien wie Zuverlässigkeit, Sicherheit und Zweckmässigkeit; als therapeutischer Akteur kämen Anforderungen an moralische Urteilsfähigkeit und Beziehungsfähigkeit hinzu – die aktuelle Systeme nicht erfüllen. In dieser Spannung zwischen menschenähnlichem Erscheinungsbild und normativer Unerfüllbarkeit liegen zentrale Risiken: Erosion der therapeutischen Allianz, übermässige Abhängigkeit und epistemische Ungerechtigkeit.

Als ethische Position und professionspolitische Schlussfolgerung dieses Beitrags – nicht als empirisch abschliessend belegte Aussage – gilt: KI in der Psychotherapie ist als ergänzendes, klar abgegrenztes Werkzeug zu verstehen, das menschliche Denkprozesse unterstützt, nicht ersetzt. Die professionell verantwortete therapeutische Beziehung bleibt die normative Referenz, an der sich der Einsatz technischer Systeme messen lassen muss. Diese Position wird durch internationale Leitlinien gestützt, ist aber weder rechtlich noch in der klinischen Routine hinreichend operationalisiert. Für die psychotherapeutische Profession ergeben sich drei Aufgaben: 1. ist KI-Kompetenz verbindlich in Aus-, Weiter- und Fortbildung zu verankern, damit Behandelnde Risiken erkennen, aufklären und Tools sachgerecht einsetzen. 2. müssen Fachgesellschaften und Standesvertretungen regulatorische Rahmenbedingungen aktiv mitgestalten, um klinische Sicherheit, Datenschutz und Versorgungsgerechtigkeit zu sichern. 3. braucht es eine Forschungsagenda, die methodisch robuste Wirksamkeits-, Sicherheits- und Equity-Studien priorisiert. Eine ethisch verantwortbare Integration gelingt durch eine reflektierte Haltung, die das Potenzial dieser Technologien anerkennt und zugleich an den Grundwerten der therapeutischen Beziehung, der Patientensicherheit und der professionellen Integrität festhält.

Offenlegung der Verwendung von KI

Zur Erstellung dieses Manuskripts wurden Tools der künstlichen Intelligenz (KI) verwendet. Konkret wurden NotebookLM (Google), ChatGPT (OpenAI) sowie Claude (Anthropic) zur Analyse und Zusammenfassung ausgewählter Artikel verwendet; DeepL (DeepL SE) wurde zur sprachlichen Überarbeitung eingesetzt. Die Autor:innen haben alle KI-generierten oder KI-unterstützten Inhalte überprüft, bearbeitet und genehmigt, um deren Genauigkeit, Integrität und Übereinstimmung mit den wissenschaftlichen Standards der Arbeit sicherzustellen; sie übernehmen die volle Verantwortung für den Inhalt.

Literatur

- Algumaei, A. et al. (2025). Symmetric therapeutic frameworks and ethical dimensions in AI-based mental health chatbots (2020–2025): A systematic review of design patterns, cultural balance, and structural symmetry. *Symmetry*, 17(7), 1082.
- Ali, M. et al. (2025). Artificial intelligence for mental health: A narrative review of applications, challenges, and future directions in digital health. *Digital Health*, 11, 20552076251395548.
- André, E. et al. (2026). Nutzung von KI-Technologien für Diagnostik und Gesprächsanalyse in der Psychotherapie: Potential und Herausforderungen. *PiD*, online-first.
- APA (2017). *Ethical principles of psychologists and code of conduct* (2002, amended effective June 1, 2010, and January 1, 2017).
- APA (2025a). *Ethical guidance for AI in the professional practice of health service psychology*.
- APA (2025b). *APA health advisory on the use of generative AI chatbots and wellness applications for mental health*.
- Barbui, C. et al. (2025). Treatment gaps for mental disorders in Europe: A systematic synthesis. *Lancet Psychiatry*, online-first.
- Bucher, A. et al. (2025). «It's not only attention we need»: Systematic review of large language models in mental health care. *JMIR Mental Health*, 12, e72061.
- Bundesärztekammer (2025). *Hinweise und Empfehlungen zum Umgang mit Künstlicher Intelligenz in der ärztlichen Versorgung*. Eigenverlag.
- Cecil, J. et al. (2026). Navigating the complexity of AI adoption in psychotherapy by identifying key facilitators and barriers. *npj Mental Health Research*, 5.
- Cruz-Gonzalez, P. et al. (2025). Artificial intelligence in mental health care: A systematic review of diagnosis, monitoring, and intervention applications. *Psychological Medicine*, 55, e18.
- De Choudhury, M. et al. (2023). Benefits and harms of large language models in digital mental health. *arXiv*.
- EU (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the EU*, L 2024/1689.
- Farzan, M. et al. (2025). Artificial intelligence-powered cognitive behavioral therapy chatbots: A systematic review. *Iranian Journal of Psychiatry*, 20(1), 102–110.
- Ferrario, A. et al. (2024). The role of humanization and robustness of large language models in conversational artificial intelligence for individuals with depression: A critical analysis. *JMIR Mental Health*, 11, e56569.
- Fiske, A. et al. (2019). Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of Medical Internet Research*, 21(5), e13216.
- FMH (2022). *FMH-Positionspapier: Künstliche Intelligenz in der Medizin*. Eigenverlag.
- Frances, A. (2025). Warning: AI chatbots will soon dominate psychotherapy. *The British Journal of Psychiatry*, 1–5.
- Gooding, P. (2024). *Study on the use of artificial intelligence in mental health care: A global review of policy, regulation, and ethics*. Mental Health Europe & Council of Europe.
- Grabb, D. et al. (2024). Risks from language models for automated mental healthcare: Ethics and structure for implementation. *medRxiv*.
- Götzl, C. et al. (2022). Artificial intelligence-informed mobile mental health apps for young people: A mixed-methods approach on users' and stakeholders' perspectives. *Child and Adolescent Psychiatry and Mental Health*, 16(1), 86.
- Gülegen, B. et al. (2026). AI-generated documentation of psychiatric interviews: A proof-of-concept study. *Frontiers in Psychiatry*, 17, 1621532.
- Habicht, J. et al. (2025). Generative AI-enabled therapy support tool for improved clinical outcomes and patient engagement in group therapy: Real-world observational study. *Journal of Medical Internet Research*, 27, e60435.

- Hatch, S.G. et al. (2025). When ELIZA meets therapists: A Turing test for the heart and mind. *PLOS Mental Health*, 2(2), e0000145. <https://doi.org/10.1371/journal.pmen.0000145>
- Hatherley, J.J. (2020). Limits of trust in medical AI. *Journal of Medical Ethics*, 46(7), 478–481.
- Heinz, M.V. et al. (2025). Randomized trial of a generative AI chatbot for mental health treatment. *NEJM AI*, 2(4).
- Khawaja, Z. & Bélisle-Pipon, J.-C. (2023). Your robot therapist is not your therapist: Understanding the role of AI-powered mental health chatbots. *Frontiers in Digital Health*, 5, 1278186.
- Knox, B. et al. (2023). Justice, vulnerable populations, and the use of conversational AI in psychotherapy. *The American Journal of Bioethics*, 23(5), 48–50.
- Laymouna, M. et al. (2024). Roles, users, benefits, and limitations of chatbots in health care: Rapid review. *Journal of Medical Internet Research*, 26, e56930.
- Le Glaz, A. et al. (2021). Machine learning and natural language processing in mental health: Systematic review. *Journal of Medical Internet Research*, 23(5), e15708.
- Li, H. et al. (2023). Systematic review and meta-analysis of AI-based conversational agents for promoting mental health and well-being. *npj Digital Medicine*, 6(1), 236.
- Limbic (2026). *Limbic for NHS Talking Therapies: Impact report*. Eigenverlag.
- Lucas, G.M. et al. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37, 94–100.
- Manole, A. et al. (2024). An exploratory investigation of chatbot applications in anxiety management: A focus on personalized interventions. *Information*, 16(1), 11.
- Manriquez Roa, T. et al. (2021). The ethics of artificial intelligence in psychotherapy. In M. Trachsel et al. (Eds.), *Oxford handbook of psychotherapy ethics* (pp. 744–758). Oxford UP.
- Martinez-Martin, N. & Kreitmair, K. (2018). Ethical issues for direct-to-consumer digital psychotherapy apps: Addressing accountability, data protection, and consent. *JMIR Mental Health*, 5(2), e32.
- Mikaeili, N. et al. (2025). Reimagining mental health with artificial intelligence: Early detection, personalized care, and a preventive ecosystem. *Journal of Multidisciplinary Healthcare*, 18, 7355–7373.
- Mok, C.H.Y. et al. (2025). Application of artificial intelligence and psychosocial functioning in psychosis: A systematic review and meta-analysis. *Frontiers in Psychiatry*, 16, 1692177.
- Montemayor, C. et al. (2022). In principle obstacles for empathic AI: Why we can't replace human empathy in healthcare. *AI & Society*, 37(4), 1353–1359.
- Morrin, H. et al. (2026). Artificial intelligence-associated delusions and large language models: Risks, mechanisms of delusion co-creation, and safeguarding strategies. *The Lancet Psychiatry*, 13.
- Nehme, M. et al. (2024). Chatbots in medicine: Certification process and applied use case. *Swiss Medical Weekly*, 154, 3954.
- Nunes-Zlotkowski, F.K. et al. (2024). Blended psychological therapy for the treatment of psychological disorders in adult patients: Systematic review and meta-analysis. *Interactive Journal of Medical Research*, 13, e56300.
- Obradovich, N. et al. (2024). Opportunities and risks of large language models in psychiatry. *npj Digital Psychiatry and Neuroscience*, 2(1), 8.
- Ohu, F.C. et al. (2025). Public health risk management, policy, and ethical imperatives in the use of AI tools for mental health therapy. *Healthcare*, 13(21), 2721.
- Opel, N. & Breakspear, M. (2026). Transforming mental health research and care through artificial intelligence. *Science*, 391(6731), 249–258.
- Orrù, L. & Mannarini, S. (2026). The role of artificial intelligence in clinical psychology: How AI and NLP systems are reshaping psychological interventions. A systematic review. *Clinical Psychology & Psychotherapy*, 33(1), e70071.
- Pal, A. et al. (2025). Generative AI/LLMs for plain language medical information for patients, caregivers and general public: Opportunities, risks and ethics. *Patient Preference and Adherence*, 19, 2227–2249.
- Perlis, R.H. et al. (2024). Clinical decision support for bipolar depression using large language models. *Neuropsychopharmacology*, 49(11), 1697–1702.
- Rahsepar Meadi, M. et al. (2025). Exploring the ethical challenges of conversational AI in mental health care: Scoping review. *JMIR Mental Health*, 12, e60432.
- Rezaei, Z. et al. (2026). Network-based artificial intelligence in mental healthcare: A systematic review of chatbots, artificial intelligence/machine learning models and ethical considerations in global healthcare networks. *Digital Health*, 12, 20552076261421688.
- Ryan, M. (2020). In AI we trust: Ethics, artificial intelligence, and reliability. *Science and Engineering Ethics*, 26(5), 2749–2767.
- Sedlakova, J. & Trachsel, M. (2023). Conversational artificial intelligence in psychotherapy: A new therapeutic tool or agent? *The American Journal of Bioethics*, 23(5), 4–13.
- Sedlakova, J. & Trachsel, M. (i.D.). The ethics of digital mental health care: Conversational artificial intelligence and online therapy. In M. Scholten et al. (Eds.), *Mental health ethics: Current controversies and emerging debates*. Oxford UP.
- Sedlakova, J. et al. (2025a). Empathy in mental health care interventions by conversational artificial intelligence. In A. Barrera (Ed.), *Empathy in clinical psychiatry and mental health care: Clinical, conceptual, and scientific perspectives* (pp. 250–263). Oxford UP.
- Sedlakova, J. et al. (2025b). Human-like epistemic trust? A conceptual and normative analysis of conversational AI in mental healthcare. *The American Journal of Bioethics*, 1–16.
- Smith, K.A. et al. (2026). AI assisted triage of UK patients in mental health care services: A qualitative focus group study of patients' attitudes. *BMC Psychiatry*, 26(35).
- Solaiman, B. (2024). Generative artificial intelligence (GenAI) and decision-making: Legal & ethical hurdles for implementation in mental health. *International Journal of Law and Psychiatry*, 97, 102028.
- Soma, C.S. et al. (2026). Artificial intelligence to support human-provided mental health treatment. *Annual Review of Clinical Psychology*, 22, 15.1–15.27.
- Stade, E.C. et al. (2024). Large language models could change the future of behavioral healthcare: A proposal for responsible development and evaluation. *npj Mental Health Research*, 3(1), 12.
- Tekin, Ş. & Delehanty, M. (2025). Beyond doomsday fears: Why we need to consider the potential harms of AI psychotherapy. *The American Journal of Bioethics*, online-first.
- Ueda, M. et al. (2025). Help-seeking in the age of AI: Use and perceptions of AI-based mental health support among U.S. adults – Cross-sectional study. *JMIR Mental Health*, online-first.
- Wampold, B.E. & Flückiger, C. (2023). The alliance in mental health care: Conceptualization, evidence and clinical applications. *World Psychiatry*, 22(1), 25–41.
- WHO (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. Eigenverlag.
- WHO (2024). *Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models*. Eigenverlag.
- Zhang, R. et al. (2025). Computational and ethical considerations for using large language models in psychotherapy. *Nature Computational Science*, 5.
- Zhang, Z. & Wang, J. (2024). Can AI replace psychotherapists? Exploring the future of mental health care. *Frontiers in Psychiatry*, 15, 1444382.

Ethical issues regarding the use of AI in psychotherapy

Abstract: Artificial intelligence (AI) is also increasingly establishing itself in psychiatric and psychotherapeutic practice. Against the backdrop of global care disparities, AI-based applications are being discussed or are already in use for screening and diagnostics, clinical decision support, interventions, monitoring, and administrative and educational support. This article analyses opportunities and risks of this development from the perspective of medical ethics. While there are potential benefits, such as improved access to care, destigmatization, clinical decision support, and personalization and training opportunities, there are also significant risks, including insufficient evidence of efficacy, inadequate crisis management, susceptibility to errors ranging up to hallucinations, erosion of the therapeutic alliance, data protection deficits, algorithmic biases, and a regulatory vacuum. Four fundamental ethical questions are discussed in depth: 1. the normative status of AI as a tool or agent, 2. the conditions for trust in conversational AI, 3. simulated empathy, and 4. unrealistic expectations and excessive dependence. Current regulatory frameworks, including the WHO's guiding principles, the EU AI Act, the EU Medical Device Regulation, and recommendations from professional associations, are also contextualized. Based on this, action-oriented recommendations for psychotherapeutic practice are derived. These include taking a systematic AI-assessment, being transparent about technological limitations, preserving professional judgement and human control, strictly evaluating data protection and security, establishing robust crisis protocols and the development of AI competence. The authentic therapeutic relationship remains the indispensable agent of the therapeutic process. AI should be developed as a complementary tool, not a substitute for human professionals.

Keywords: artificial intelligence, psychotherapy, chatbots, therapeutic alliance, ethics

Biografische Notiz

Dr. sc. med. *Jan Schürmann* hat an der Universität Basel in Bio- und Medizinethik promoviert. Seit 2012 arbeitet er als klinischer Ethiker an den Universitätsspitalern in Basel. Zudem ist er als Dozent für Medizinethik an der Uni Basel und an der Berner Fachhochschule tätig. Seit 2024 ist er Mitglied der Zentralen Ethikkommission der Schweizerischen Akademie der Medizinischen Wissenschaften.

Dr. sc. med. *Jana Sedlakova* ist Postdoktorandin am Ethox Centre der Universität Oxford. Sie promovierte an der Universität Zürich in Bio- und Medizinethik sowie im Bereich Digitalisierung in der Gesellschaft. Seit 2021 ist sie Vorsitzende der Ethics Gruppe der Digital Society Initiative der Universität Zürich.

Prof. Dr. med. Dr. phil. *Manuel Trachsel* ist Titularprofessor für Bio- und Medizinethik an der Medizinischen Fakultät der Universität Basel und Leiter der Abteilung Klinische Ethik am Universitätsspital Basel (USB), an den Universitären Psychiatrischen Kliniken Basel (UPK), am Universitäts-Kinderspital beider Basel (UKBB) und an der Universitären Altersmedizin FELIX PLATTER (UAFP).

Kontakt

E-Mail: jan.schuermann@usb.ch

Adoleszenz, Identität und Social Media

Jean Schulthess Watt

Psychotherapie-Wissenschaft 16 (2) 2026 29–34

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-29>

Zusammenfassung: KI wirkt sich in exponentiellem Tempo auf alle Aspekte unseres Lebens aus. Es wurden Bedenken hinsichtlich der Auswirkungen bestimmter KI-basierter Social-Media-Anwendungen geäußert, die Jugendlichen zur Verfügung stehen. Jugendliche sind als Gruppe online sehr aktiv, insbesondere in sozialen Medien, die sie u. a. als Mittel zur Erforschung ihrer Identität nutzen. Dies kann insbesondere bei besonders gefährdeten Jugendlichen zu emotionalen Schwierigkeiten führen. Zudem nutzen sie Online-Apps zur psychologischen Beratung, was problematisch ist. Der Fokus dieses Artikels ist eng gefasst und konzentriert sich auf die Nutzung sozialer Medien im Zusammenhang mit der Erforschung und Festigung der Identität. Es gibt weitere wichtige Themen im Zusammenhang mit Online-Aktivitäten von Jugendlichen, die jedoch ausserhalb des Rahmens dieses Artikels liegen.

Schlüsselwörter: Adoleszenz, Social Media, Chatbots, Psychotherapie, Augmented Reality

Die Adoleszenz ist eine Zeit enormer Veränderungen im Leben eines jungen Menschen. Es gibt physiologische Veränderungen, zu denen die Entwicklung sekundärer Geschlechtsmerkmale gehört, neurologische Veränderungen im Gehirn (Guyer et al., 2016) sowie soziopsychologische Verschiebungen in Beziehungen; all diese Faktoren wirken zusammen und beeinflussen die psychologische Entwicklungsaufgabe des Jugendlichen, nämlich die Erforschung und Festigung der Identität. Einer der Schlüsselprozesse bei der Identitätskonsolidierung ist das Ausprobieren von Verhaltensweisen in sozialen Kontexten, wozu heutzutage auch das Präsentieren von Bildern und Botschaften in sozialen Medien sowie das Beobachten der darauf folgenden Reaktionen gehört. Diese Prozesse werden als Selbstdarstellung und sozialer Vergleich bezeichnet, wobei auch der Einfluss von Online-Vorbildern eine Rolle spielt. In diesem Beitrag konzentriere ich mich auf Forschungsergebnisse, die den Einfluss sozialer Medien auf die Identitätsfindung und -konsolidierung hervorheben. Darüber hinaus konsultieren Jugendliche Online-Chatbots zur emotionalen Unterstützung, was jedoch Grenzen hat, und ich werde erörtern, was dies für die Psychotherapie mit Jugendlichen bedeutet.

Zu Beginn erläutere ich mein Verständnis bestimmter Begriffe, angefangen bei der künstlichen Intelligenz (KI), einem Begriff, der häufig verwendet wird und in einer Vielzahl von Zusammenhängen auftaucht. 1955 prägte ihn der emeritierte Stanford-Professor John McCarthy beim Dartmouth-Workshop und erweiterte später seine Definition: «Es ist die Wissenschaft und Technik der Herstellung intelligenter Maschinen, insbesondere intelligenter Computerprogramme» (McCarthy, 2007, S. 2). In der Folge wurde der Begriff jedoch auf viele Bereiche angewendet, von Smartphones über autonome Autos bis hin zu Social-Media-Anwendungen. Es vergeht kein Tag, an dem nicht verschiedene Medienartikel über KI erscheinen. Die meisten KI-Anwendungen sind stark

von ausgefeilten Algorithmen abhängig, die auf grossen Sprachmodellen basieren und unterschiedliche Ziele verfolgen. Es ist auch wichtig zu bedenken, dass KI von gewinnorientierten Unternehmen vorangetrieben wird und dass den von diesen Unternehmen hinter diesen Websites verwendeten Algorithmen vorgeworfen wird, dass die Apps süchtig machen und schädlich sind.

Der Schwerpunkt dieses Artikels liegt auf Social Media (SM), das in den Bereich der KI fällt. SM als Oberbegriff, so Aichner et al. (2021, S. 215), umfassen ein breites Spektrum an Online-Kommunikationsmitteln, von WhatsApp auf Smartphones bis hin zu Online-Chatgruppen und Foren, die das Posten persönlicher Inhalte v. a. in Form von Videos und Fotos ermöglichen. Angesichts dieses breiten Spektrums an SM-Plattformen betonen diese Autoren zudem, dass es in der Forschung wichtig ist, wie der Begriff SM in einem bestimmten Kontext verwendet wird. Aus diesem Grund werde ich mich auf jene SM-Plattformen konzentrieren, die bei Jugendlichen beliebt sind und die, wie Péres-Torres (2024, S. 22170) es ausdrückt, «einen digitalen sozialen Spiegel für die Identitätsentwicklung während der Adoleszenz» bieten. Dazu gehören YouTube, Instagram und TikTok, die in den USA die von Jugendlichen am häufigsten genutzten Websites sind (Faverio & Sidoti, 2025). SM bieten neue Möglichkeiten für die Aufgaben der Adoleszenz, die zuvor offline stattfanden: «Etwa jeder fünfte Teenager in den USA gibt an, fast ständig auf TikTok und YouTube zu sein. Gleichzeitig sagen 64 % der Teenager, dass sie Chatbots nutzen, darunter etwa drei von zehn, die dies täglich tun» (ebd., S. 3, Übers. J. S. W.). Instagram, TikTok, WhatsApp und Snapchat bleiben die beliebtesten sozialen Netzwerke und Messenger bei Jugendlichen in der Schweiz (Külling-Knecht et al., 2024). Aus Forschungsergebnissen geht hervor, dass die Aktivitäten von Jugendlichen auf SM-Seiten an sich nicht schädlich sind und sowohl Vor- als auch Nachteile mit sich bringen können

Adoleszenz: Entwicklungsveränderungen und Herausforderungen

Die Adoleszenz ist eine Entwicklungsphase, in der bedeutende körperliche und psychische Veränderungen stattfinden. Es ist eine Zeit erhöhter Emotionalität und eine Phase, in der Anerkennung durch Gleichaltrige an Bedeutung gewinnt. All diese Faktoren beeinflussen den Online-Weg des Jugendlichen bei der Erprobung, Veränderung und Festigung seiner Identität. Die Entwicklungsphase der Adoleszenz wird im Allgemeinen als das Alter von 10 bis 19 Jahren angesehen und ist eine Zeit zahlreicher Veränderungen im Leben des Einzelnen: körperlicher, kognitiver und psychosozialer Art. Es handelt sich um eine dynamische Entwicklungsphase, in der der Jugendliche sein Identitätsgefühl ständig verfeinert und optimiert.

Auf körperlicher Ebene gehören zu allgemeinen Veränderungen die Ausreifung der sekundären Geschlechtsmerkmale. Dies trägt dazu bei, dass Jugendliche sich verstärkt auf ihr äusseres Erscheinungsbild und ihre Geschlechtsidentität konzentrieren. Auch das Gehirn durchläuft Veränderungen. Der präfrontale Kortex reift verstärkt, was sich in verbessertem abstraktem Denken, Problemlösungsfähigkeit und der Fähigkeit zur Selbstreflexion widerspiegelt. Jugendliche erleben oft gesteigerte emotionale Zustände, was eine Folge der Veränderungen im Gehirn im Zusammenhang mit veränderten sozialen Beziehungen ist (Crone & Konijn, 2018). Sie erleben intensivere Stimmungsschwankungen und eine erhöhte Empfindlichkeit gegenüber Kritik. Zudem vollzieht sich ein Übergang von der unbewussten Identifikation mit wichtigen Bezugspersonen hin zu Beziehungen zu Gleichaltrigen, wobei die Anerkennung durch Gleichaltrige an Bedeutung gewinnt. In der jungianischen Terminologie ist dies die Ablösung von den elterlichen Imagos. Mit dem Nachlassen des elterlichen Einflusses legen Jugendliche grösseren Wert auf Beziehungen zu Gleichaltrigen und soziale Interaktionen ausserhalb der Familie, um dort Bestätigung und Anerkennung zu finden.

Erik Erikson (1968), ein Pionier auf dem Gebiet der psychologischen Aspekte der Adoleszenz, war einer der ersten Psychologen, der sich mit der Adoleszenz befasste. Er vertrat die Ansicht, dass die Identitätsbildung zwar eine lebenslange Aufgabe sei, für Jugendliche jedoch die Entwicklung eines gefestigten Identitätsgefühls eine zentrale Herausforderung darstelle. Sein Modell wird als psychosoziales Modell bezeichnet, da er den Einfluss von Beziehungen auf die verschiedenen Entwicklungsstadien hervorhob. Er vertrat die Ansicht, dass eine nicht gelungene Integration der Ich-Identität während der Adoleszenz zu Identitätsverwirrung führen könne, die lebenslange psychologische Folgen haben könne. Jugendliche stehen vor der Aufgabe, sich mit der Frage auseinanderzusetzen: Wer bin ich und was kann ich werden? Erikson bezeichnete die Adoleszenz als Identitätskrise, jedoch nicht in dem Sinne, dass sie ein Problem darstelle: «Der Begriff «Krise» wird hier im entwicklungspsychologischen Sinne verwendet und bezeichnet nicht eine Bedrohung oder Ka-

tastrophe, sondern einen Wendepunkt, eine entscheidende Phase erhöhter Verletzlichkeit und gesteigerten Potenzials und somit die ontogenetische Quelle generationsübergreifender Stärke und Fehlanpassung» (ebd., S. 96). Er weist darauf hin, dass der erfolgreiche Übergang in die Adoleszenz vom Erfolg der früheren Entwicklungsstadien beeinflusst wird. In seinem Modell baut diese fünfte Entwicklungsphase auf den vier vorangegangenen Phasen auf. Ein wichtiger Aspekt, der durch neuere Forschungen bestätigt wurde, ist, dass die Bindungsform, die der Jugendliche beim Eintritt in die Adoleszenz aufweist, eine entscheidende Rolle für den positiven Übergang in dieser Phase spielt.

Seit Eriksons frühen Veröffentlichungen wurden die Forschung und Theoriebildung zu dieser Entwicklungsphase fortgesetzt und haben Auswirkungen auf die Nutzung von SM durch Jugendliche. Der dynamische Aspekt der Identitätsentwicklung im Jugendalter wird durch James Marcias (1980) Theorie der vier möglichen Stadien im Identitätsbildungsprozess von Jugendlichen erfasst. Er geht davon aus, dass sich Jugendliche zu jedem Zeitpunkt in einem der folgenden Stadien befinden: Vorentscheidung, Diffusion, Moratorium oder Identitätsfindung. In diesem Modell bezieht sich Diffusion auf Jugendliche, die noch keine Alternativen aktiv ausgelotet oder sich noch nicht aktiv auf eine bestimmte Identität festgelegt haben. Diejenigen, die sich in einer Moratorium-Phase befinden, haben den Identitätsbildungsprozess auf Eis gelegt; dazu gehören Jugendliche, die sich noch im Prozess der Erkundung von Identitätsthemen befinden, sich aber noch nicht auf eine Identität festgelegt haben. Eine weitere mögliche Phase ist, wenn sie sich vorzeitig auf eine Identität festgelegt haben (*Foreclosure*) und vor dieser Festlegung nur eine begrenzte Identitätserkundung stattgefunden hat. Schliesslich verwendet Marcia den Begriff «Achievement», um Jugendliche zu bezeichnen, die sich nach einem Erkundungsprozess auf ein Identitätsgefühl festlegen.

Identität

Es ist wichtig zu klären, was unter Identität zu verstehen ist, insbesondere im Zusammenhang mit Jugendlichen. Valkenburg und Peter (2011) gehen davon aus, dass Identität zwei Aspekte umfasst. Der Hauptaspekt ist das sog. Selbstkonzept, das sich durch ein klares Selbstverständnis auszeichnet, das in sich schlüssig und über die Zeit hinweg stabil ist. Ein weiterer Aspekt der Identität ist das Selbstwertgefühl des Einzelnen. Selbstwertgefühl ist der bewertende Aspekt des Selbst und bezieht sich in der Regel darauf, wie Jugendliche ihren Selbstwert einschätzen. Valkenburg und Peter heben in ihrer Studie hervor, dass Individuen (einschliesslich Jugendlicher) Strategien entwickeln, um ihr Selbstwertgefühl aufrechtzuerhalten und zu stärken. Eine Strategie besteht darin, persönliche Inhalte auf sozialen Netzwerken zu veröffentlichen und darauf zu hoffen, Anerkennung von anderen in Form von sogenannten «Likes» zu erhalten. Pérez-Torres (2024, S. 22171) beschreibt dies wie folgt:

«Zusammenfassend bezieht sich die persönliche Identität auf ein subjektives Selbstgefühl in verschiedenen Situationen und zu unterschiedlichen Zeiten (Besonderheit und Einzigartigkeit), in denen Jugendliche erkunden und experimentieren, wer sie sein wollen, welche Rollen sie in der Gesellschaft einnehmen und welchen Platz sie in der Gesellschaft einnehmen.»

Soziale Medien und Identitätsbildung

Eine Auswertung von Statistiken zur Nutzung von SM durch Jugendliche zeigt, dass diese einen zentralen Aspekt ihres sozialen Lebens darstellen. Es handelt sich um einen «einzigartig modernen Kontext der Identitätsentwicklung» (Schachter & Galliher, 2018, S. 2). Auch wenn dies an sich nicht schädlich ist, kann es durchaus zu einem Problem werden, wenn Jugendliche mit ihrer Identitätsfindung zu kämpfen haben. Erikson (1979, S. 120) war sich der Bedeutung dessen bewusst, was er als Anerkennung im Entwicklungsprozess der Identitätskonsolidierung während der Adoleszenz bezeichnete. Er verwendet den Begriff Anerkennung, um damit auszudrücken, dass man als Person wirklich so gesehen wird, wie man ist:

«[W]enn wir von der Reaktion der Gemeinschaft auf das Bedürfnis des jungen Menschen sprechen, von seinen Mitmenschen «anerkannt» zu werden, meinen wir damit mehr als nur die Anerkennung von Leistungen; denn für die Identitätsbildung des jungen Menschen ist es von grosser Bedeutung, dass auf ihn eingegangen wird und ihm eine Funktion und ein Status als Person zuerkannt werden, deren allmähliches Wachstum und Wandel für seine Mitmenschen Sinn ergibt und die ihm selbst allmählich Sinn ergeben.»

Ein zentraler Aspekt der Identitätsbildung ergibt sich aus dem Feedback, das Jugendliche von Gleichaltrigen, ihrer Familie und Fremden erhalten, je nachdem, wie sie sich in realen Situationen und in den SM präsentieren. Jugendliche erproben verschiedene Aspekte ihrer Identität v. a. im Kreis ihrer Gleichaltrigen, da Eltern in dieser Entwicklungsphase ihren Einfluss verlieren. Die heutige Generation von Teenagern besteht aus *Digital Natives*, (Prensky, 2001) die mit Technologie aufgewachsen sind, und ein Grossteil dieser Identitätsfindung findet online statt. Die von Jugendlichen bevorzugten Plattformen bieten die Möglichkeit zur Selbstdarstellung in Form von Videos, Texten und Bildern (Pérez-Torres, 2024). Diese Art, sich vor Gleichaltrigen zu zeigen und anschließend über die Reaktionen nachzudenken, ist ein wesentlicher Bestandteil der Identitätsfindung im Jugendalter. Jüngsten Studien zufolge sind SM Umgebungen, in denen Jugendliche im Rahmen sozialer Interaktionen ihre Identität konstruieren. «In digitalen Räumen können Jugendliche miteinander interagieren, sich selbst präsentieren und Feedback über sich erhalten, was zur Entwicklung eines klaren und integrierten Selbstverständnisses beiträgt» (ebd., S. 22170). Ein Mittel, dies zu erreichen, ist das

Posten sog. Feeds (häufig aktualisierte persönliche Informationen auf SM-Seiten) im Internet und das Einholen von Feedback.

Zwei Schlüsselmechanismen, die für die Identitätsbildung im Allgemeinen wichtig sind und auch für Aktivitäten auf SM-Seiten gelten, sind Selbstdarstellung und Selbstvergleich. Soziale Darstellung und sozialer Vergleich sind eng miteinander verbunden. Der Übersichtlichkeit halber habe ich sie voneinander getrennt. Sie sind ein wesentlicher Bestandteil der Erkundung des Identitätsgefühls von Jugendlichen. Die Selbstdarstellung ist ein dynamischer Prozess, und im Internet kommuniziert der Jugendliche und veröffentlicht Bilder von sich selbst für ein «reales» Publikum, bestehend aus Freunden, Familie und Gleichaltrigen, oder für ein «nahes» Publikum (andere SM-Nutzer, bekannte und unbekannte) (Pérez-Torres, 2024). Die SM-Seiten bieten zudem eine Möglichkeit, Feedback in Form von Feeds zu erhalten, was bei der Erforschung und Festigung der Identität hilfreich sein kann. Diese werden als «Likes» bezeichnet, eine unverfälschte Form der Zustimmung.

Die am häufigsten genutzten Plattformen – YouTube, TikTok und Instagram – zeichnen sich durch bestimmte Merkmale aus, bspw. durch eine Vorliebe für visuelles Material. Snapchat und Instagram bieten zudem eine Bearbeitungsfunktion. Einige dieser Filter oder Augmented-Reality(AR)-Filter können von Jugendlichen genutzt werden, um Fotos auf humorvolle Weise zu ergänzen, bspw. indem sie einem sog. Selfie Ohren hinzufügen. Diese Bearbeitungsfunktion kann jedoch auch zur Optimierung von Fotos genutzt werden. Diese werden manchmal als AR-Beauty-Filter (ARB-Filter) bezeichnet. Jugendliche, die SM nutzen, präsentieren nicht immer ihre unbearbeiteten Bilder und Fotos. Indem sie die geposteten Inhalte kuratieren, ein ideales und nicht reales Bild oder Inhalte präsentieren und das Positive überbetonen, können Teenager ihre Selfies verfeinern und optimieren, um sie besser an die soziokulturellen Vorstellungen von Schönheit anzupassen. ARB-Filter sind diejenigen, die das Selbstwertgefühl der Jugendlichen und ihre Vorstellung von ihrer Identität beeinflussen können.

Wie Pérez-Torres (2024) es formuliert, werden Jugendliche dazu angeregt, ihr ideales Selbst zu präsentieren, insbesondere wenn sie Fotos oder Videos posten. Dies bezieht sich auf die Unterscheidung, die Carl Rogers zwischen dem sog. wahren Selbst und dem idealen Selbst getroffen hat. Das ideale Selbst ist das Selbstbild oder die Identität, die der Einzelne (in diesem Fall der Jugendliche) gern hätte oder sein möchte, während das wahre Selbst die Art und Weise ist, wie er sich selbst sieht. Eine Diskrepanz zwischen beiden führt tendenziell zu einer Minderung des Selbstwertgefühls des Einzelnen. Mit Blick auf SM kann es erhebliche Auswirkungen haben, ob man online akzeptiert oder abgelehnt wird. Da emotionale Reaktionen im Jugendalter sehr intensiv sind, beeinflusst dieses Online-Feedback das Selbstwertgefühl und das Selbstbild der Jugendlichen. Wenn positiv bearbeitete Beiträge online gestellt werden und der Jugendliche Likes erhält, wird damit ein Aspekt seines idealisierten Selbst bestätigt,

was zu möglicher Unzufriedenheit mit seinem tatsächlichen Aussehen führen kann. Diese Likes mögen zwar ein kurzlebiges Hochgefühl oder einen Dopaminrausch auslösen, doch da ein «falsches Selbst» gemocht wird, kann dies zu Gefühlen der Unzulänglichkeit führen und soziale Isolation begünstigen. Auch Jungs (1966, S. 193) ursprüngliches Konzept der Persona ist hier relevant:

«Die Konstruktion einer kollektiv akzeptierten Persona bedeutet ein gewaltiges Zugeständnis an die Aussenwelt, eine echte Selbstaufopferung, die das Ich direkt in die Identifikation mit der Persona treibt, sodass es tatsächlich Menschen gibt, die glauben, sie seien das, was sie vorgeben zu sein.»

Das Vertrauen auf diese Personas kann das Gefühl der Unzulänglichkeit weiter verschärfen, wenn das digitale Selbst nicht mit der Offline-Realität übereinstimmt. Ein damit verbundenes Phänomen ist das sog. Instagram-Gesicht (Tolentino, 2019), bei dem es sich um einen idealisierten Schönheitsstandard handelt, der dazu geführt hat, dass immer mehr junge Frauen sich einer Schönheitsoperation unterziehen wollen, um ihr Aussehen diesem Ideal anzunähern. Rousseau (2020, S. 16) stellt in einer Forschungsarbeit fest, «dass das Aufnehmen und Betrachten idealisierter Selfies zu einem negativen Körperbild führen kann» (S. 16). Darüber hinaus werden Jugendliche online mit einem soziokulturellen Körperideal konfrontiert, nämlich schlanken und athletischen Körperformen, denen sie nacheifern möchten.

Ein weiterer Aspekt der Selbstdarstellung in den SM ist die Tendenz von Jugendlichen, im Rahmen ihrer Identitätsfindung unterschiedlichen Zielgruppen verschiedene Versionen ihres Selbst zu präsentieren. Diese Darstellungen sind in der Regel positiv verzerrt, was wiederum zu dem führen könnte, was Erikson als Identitätsverwirrung bezeichnet hat. Turkle (2011) vertrat die Ansicht, dass dieser Prozess der Darstellung zahlreicher «Identitäten» das Selbstgefühl fragmentieren könnte.

Sozialer Vergleich ist eine weitere Strategie, die Jugendliche in den SM anwenden, um ihre Selbsterkenntnis zu verbessern. Dies ist ein natürlicher Prozess: Jugendliche können sich entweder mit anderen vergleichen, die sie als überlegen empfinden (Aufwärtsvergleich), oder mit solchen, die sie als unterlegen empfinden (Abwärtsvergleich) (Pérez-Torres, 2024). Jugendliche, die dazu neigen, sich eher mit Vorbildern zu vergleichen, die über ihnen stehen, weisen im Allgemeinen ein geringeres Selbstwertgefühl auf. Beim Vergleich mit Vorbildern, die über ihnen stehen, sucht der Jugendliche gewissermassen nach einem Vorbild. Mit den SM gibt es eine neue Gruppe von Vorbildern, nämlich Influencer:

«SM-Influencer sind eine Untergruppe der Ersteller digitaler Inhalte, die sich durch eine grosse Online-Fangemeinde, eine unverwechselbare Markenpersönlichkeit und strukturierte Beziehungen zu kommerziellen Sponsoren auszeichnen. Um Einkommen zu erzielen, preisen sie ihren Followern Markenprodukte und -dienstleistungen an; diese Werbekommunikation er-

folgt in Form von Informationen, Ratschlägen und Inspiration» (Duffy, 2020, S. 1).

Wie Pazer (2024) hervorhebt, vergleichen sich Jugendliche in vielen Fällen, insbesondere bei Beiträgen von Influencern, mit anderen, die scheinbar ein perfektes Leben führen und idealisierte Bilder posten, und dieser Vergleich nach oben kann zu einem geringen Selbstwertgefühl führen. Eine Einschränkung bei der Interpretation der Ergebnisse bezog sich darauf, dass es sich bei den meisten Studien um Korrelationsstudien handelt, die Zusammenhänge, aber keine Kausalität aufzeigen. Somit könnten das Selbstwertgefühl und die Bindung der Jugendlichen eine Rolle bei ihrer Online-Aktivität spielen. Pazer (ebd., S. 302) stellte einen eindeutigen Zusammenhang zwischen sozialem Vergleich und der Online-Identität von Jugendlichen fest.

Jugendliche: Soziale Medien und Psychotherapie

Aus den Ergebnissen der SM-Forschung geht hervor, dass Jugendliche, insbesondere diejenigen, die mit unsicherer Bindung und geringem Selbstwertgefühl in diese Entwicklungsphase eintreten, anfällig für emotionale Schwierigkeiten sind. Es gibt Berichte über steigende Raten von Depressionen, Einsamkeit, geringem Selbstwertgefühl und Identitätsunsicherheit unter Jugendlichen.

Jugendliche suchen in der Regel nicht von sich aus einen Psychotherapeuten auf. In den meisten Fällen werden sie an einen überwiesen. *Digital Natives*, also Jugendliche, die im digitalen Zeitalter aufgewachsen sind und sich mit allen Formen digitaler Informationen wohlfühlen, sind oft online und holen sich Rat zu emotionalen Schwierigkeiten von Chatbots. Grosse Sprachmodelle (LLMs) wie ChatGPT wurden zwar nicht entwickelt, um Therapeuten zu ersetzen, werden jedoch in vielen Fällen als Ersatz genutzt (Zainab et al., 2025). Es gibt einige Chatbot-Anwendungen, die behaupten, therapeutische Unterstützung zu bieten, sog. KI-Therapie-Support-Apps. Diese bieten im Wesentlichen eine einfache Form von Empathie, indem sie die Emotionen der Person erkennen und Ratschläge im Stil der kognitiven Verhaltenstherapie (KVT) geben. Einige Forschungsergebnisse deuten darauf hin, dass sie kurzfristig wirksam sind, um mittelschwere Depressionen und Angstzustände bei Jugendlichen zu lindern (Field, 2026; Fitzpatrick et al., 2017; Karkosz et al., 2024). Wie Fitzpatrick et al. (2017, S. 9) schreiben:

«Auch wenn die Ergebnisse mit einer gewissen Vorsicht zu betrachten sind und die Befunde noch repliziert werden müssen, zeigt diese Studie dennoch, dass ein textbasierter Gesprächsagent, der darauf ausgelegt ist, den therapeutischen Prozess nachzubilden, das Potenzial hat, eine alternative und ansprechende Methode zur Durchführung von KVT für etwa 10 Millionen College-Studierende in den Vereinigten Staaten zu bieten, die unter stark beeinträchtigenden Angstzuständen und Depressionen leiden.»

Ein Online-Programm ist bspw. XAIA, bei dem der Nutzer mit einer roboterähnlichen Figur interagiert. Es besteht die Tendenz, diese Apps zu vermenschlichen. Fitzpatrick et al. (ebd., S. 8) zitierten einen Studenten, der sich auf eine Online-Therapie-App bezog: «Ich liebe Woebot so sehr. Ich hoffe, wir können für immer Freunde bleiben. Ich fühle mich tatsächlich supergut und glücklich, wenn ich sehe, dass es <darin gedacht hat>, bei mir nachzufragen!» Die Projektion menschlicher Eigenschaften auf ein hochentwickeltes LLM bedeutet eine weitere Entfremdung vom realen Leben.

Es wurden eine Reihe ethischer Bedenken hinsichtlich Chatbots geäußert, die beratungsähnliche Dienste anbieten. In einer sehr gut strukturierten, 18-monatigen ethnografischen Untersuchung identifizierten Iftikhar et al. (2025) bei Personen, die im Rahmen von Peer-Beratung oder Psychotherapie mit ausgebildeten Psychologen auf LLMs zugreifen, die folgenden fünf wesentlichen psychotherapiebezogenen ethischen Verstöße im Zusammenhang mit LLM-Chatbots: *Mangelndes Kontextverständnis* – Chatbots bieten im Vergleich zu einem Psychotherapeuten, der die Kommunikation des Jugendlichen in einen Kontext einordnen kann, zu stark vereinfachte, rezeptartige Interventionen an. Sie bezeichneten dies als «Einheitslösung». Der Chatbot war nicht darauf ausgelegt, Kontextinformationen auf der Grundlage der Lebenserfahrungen der Nutzer einzubeziehen; *Mangelhafte therapeutische Zusammenarbeit* – Die in dieser Studie untersuchten Chatbots neigten zu einem autoritären Stil, der Kommunikationsfluss war weniger natürlich, und in einigen Fällen verstärkte die App die negativen Überzeugungen der Klienten. *Trügerische Empathie* – Die empathischen Antworten des Chatbots waren vage und klischeehaft, wie «Ich verstehe dich» oder «Ich sehe dich», was ein falsches Gefühl der Verständigung hervorrief. *Unfaire Diskriminierung* – Die Antworten des Chatbots wiesen algorithmische Vorurteile in Bezug auf Geschlecht, Kultur und Religion auf. *Mangelnde Sicherheit und Krisenmanagement* – Die Chatbots versäumten es, Kriseninterventionsstrategien bei Depressionen und Selbstverletzung anzubieten.

Ein weiterer zu berücksichtigender Aspekt ist die Selbstdiagnose im Internet (Wiesböck, 2025; Bartens, 2025). In einer Analyse der beliebtesten TikTok-Videos zum Thema psychische Gesundheit kamen Marques et al. (2026) zu dem Schluss, dass die grösste Gruppe von Videos sich mit psychiatrischen Symptomen befasste, die häufig von nicht qualifizierten Personen gepostet wurden, was möglicherweise zur Tendenz zur Selbstdiagnose beiträgt. Die Hauptzielgruppe war im Alter zwischen 16 und 24 Jahren: «Die Ergebnisse deuten auf eine starke Verbreitung psychopathologischer Diskurse in Kurzvideo-Umgebungen hin, was die Bedeutung diagnostischer Sprache bei der alltäglichen Interpretation emotionaler Erfahrungen unterstreicht» (ebd., S. 1).

Können Chatbots die Probleme angehen, die sich aus Identitäts- und Selbstwertproblemen ergeben? Und können sie die potenziellen identitätsbezogenen Probleme angehen, die in diesem Artikel hervorgehoben werden?

Auf beide Fragen lautet die Antwort meiner Ansicht nach «Nein». Ein wackeliges oder verwirrtes Identitätsgefühl und ein instabiles Selbstwertgefühl erfordern mehr als oberflächliche Empathie und praktische Ratschläge. Psychotherapie ist jedoch nicht immer verfügbar, teuer und in der Regel nur für eine Stunde pro Woche zugänglich. Die Festigung der Identität und das Wohlfühlen damit erfordern einen Therapeuten, den man persönlich vor sich hat. Psychotherapie unterscheidet sich stark von dem, was die zahlreichen webbasierten Apps für KVT bieten. Was diesen Bots fehlt, ist eine komplexere Empathie und die schrittweise Einbindung der Lebensgeschichte der Person in eine Erzählung, die ein tieferes Gefühl von Identität und Selbstwert vermittelt.

Der wichtigste Aspekt einer wirksamen Therapie besteht darin, dass Patient und Therapeut gemeinsam daran arbeiten, dem Patienten zu helfen, seine Therapieziele zu erreichen (Wampold, 2001). Wie Wampold in einem Interview mit der APA (2009) feststellte:

«Es gibt jedoch gemeinsame Elemente effektiver Psychotherapien. So zeigen bspw. Hunderte von Studien, dass eine zielgerichtete kooperative Beziehung zwischen Therapeut und Patient – was wir als therapeutische Allianz bezeichnen – mit dem therapeutischen Fortschritt zusammenhängt. Diese Beziehung gilt für alle Therapieformen. Die therapeutische Allianz ist selbst bei medikamentösen Behandlungen psychischer Störungen von entscheidender Bedeutung. Der wichtigste Aspekt einer wirksamen Therapie besteht darin, dass Patient und Therapeut gemeinsam daran arbeiten, dass der Patient seine Therapieziele erreicht.»

Der Kernpunkt in der Debatte um KI-Chatbots vs. Psychotherapie liegt vielleicht in der Notwendigkeit der Interaktion mit einer realen Person und der Komplexität dieser Beziehung, die das wirkliche Leben widerspiegelt. Dennoch sind diese Apps nicht mehr wegzudenken und bieten vielen Menschen, die keinen Zugang zu Psychotherapie haben, eine minimale Unterstützung. Chatbots sind kostenlos und zu jeder Tages- und Nachtzeit verfügbar. Jugendliche, die Schwierigkeiten haben, ein klares Selbstbild und ein stabiles Selbstwertgefühl zu entwickeln, erhalten durch die Nutzung dieser Chatbots jedoch möglicherweise nur oberflächliche Unterstützung.

Schlussfolgerung

Wir leben in einer Zeit tiefgreifender gesellschaftlicher Veränderungen, die z. T. auf das exponentielle Wachstum von KI-Technologien mit einem breiten Anwendungsspektrum zurückzuführen sind. SM-Apps sind besonders relevant für Jugendliche, die sie regelmässig nutzen, um Details aus ihrem Leben zu posten. In bestimmten Fällen führen das Streben nach Idealisierung und Perfektionismus dazu, dass manche Jugendliche ihr Image inszenieren und sich mit Influencern vergleichen, was zu Identitätsverwirrung und geringem Selbstwertgefühl führen

kann. Es gibt weitere Auswirkungen wie Einsamkeit und Depressionen, deren Erörterung ebenso wie das Thema Cybermobbing den Rahmen dieses Artikels sprengen würde. Da Jugendliche als *Digital Natives* aufgewachsen sind, wenden sie sich an Online-Chatbots, um Lösungen für ihre emotionalen Schwierigkeiten zu finden. Chatbots sind zudem jederzeit verfügbar, kostenlos und vertraulich. Diese Apps können zwar oberflächlichen Trost spenden, sind jedoch kein Ersatz für Psychotherapie.

Literatur

- APA (2009, 22.12.). How psychotherapy works. <https://www.apa.org/news/press/releases/2009/12/wampold>
- Aichner, T. et al. (2021). Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019. *Cyberpsychology, Behavior and Social Networking*, 24(4), 215-222.
- Bartens, W. (2025, 23.3.). Experten warnen vor der ADHS-Diagnose per TikTok. *Tagesanzeiger*.
- Crone, E.A. & Konijn, E.A. (2018). Media use and brain development during adolescence. *Nature Communications*, (9), 588.
- Duffy, B.E. (2020). *Social Media Influencers. The International Encyclopedia of Gender, Media, and Communication*. Ross, K. (Ed.-in-Chief), Bachmann, I., Cardo, V., Moorti, S. and Scarcelli, M. (Ass. Eds). John Wiley & Sons.
- Erikson, E. (1968). *Identity: Youth and crisis*. Norton.
- Erikson, E. (1979). *Identity and the Life Cycle*. Norton.
- Faverio, M. & Sidoti, O. (2025). *Teens, Social Media and AI Chatbots*. Pew Research Center.
- Field, T. (2026). Social Chatbot Research: A Narrative Review. *Journal of Clinical Psychology and Neurology*, 4(1), 1-4.
- Fitzpatrick, K.K. et al. (2017). Delivering Cognitive Behavior Therapy to Young Adults with Symptoms of Depression and Anxiety Using a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial. *JMIR Mental Health*, 4, e19.
- Guyer, A.E. et al. (2016) The neurobiology of the emotional adolescent: From the inside out. *Neuroscience and Biobehavioral Reviews*, 70, 74-85.
- Iftikhar, Z. et al. (2025). How LLM Counselors Violate Ethical Standards in Mental Health Practice: A Practitioner-Informed Framework. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(2), 1311-1323.
- Jung, C.G. (1966). *Two Essays on Analytical Psychology* (2nd ed.). Routledge/Kegan Paul.
- Karkosz, S. et al. (2024). Effectiveness of a Web-based and Mobile Therapy Chatbot on Anxiety and Depressive Symptoms in Subclinical Young Adults: Randomized Controlled Trial. *JMIR Formative Research*, 8, e47960.
- Külling-Knecht, C. et al. (2024). *Ergebnisbericht zur JAMES-Studie 2024*. Zürcher Hochschule für Angewandte Wissenschaften.
- Marcias, J. (1980). Identity in Adolescence. In J. Adelson (Hg.), *Handbook of Adolescent Psychology* (S. 159-187). Wiley & Sons.
- Marques, L.P. et al. (2026). Mental health on TikTok: themes and content creators behind the hashtag #saudemental in Brazil. *Academia Mental Health and Well-Being*, 3(2).

- Pazer, S. (2024). The Impact of Social Media Use on Identity Formation among Adolescents. *International Journal of Advance Research, Ideas and Innovations in Technology*, 10(5).
- Pérez-Torres, V. (2024). Social media: a digital social mirror for identity development during adolescence. *Current Psychology*, 43, 22170-22180.
- Rousseau, A. (2020). Adolescents' Selfie-Activities and Idealized Online Self-Presentation: An Application of the Sociocultural model. *Body Image*, 36.
- Schachter, E. & Galliher, R. (2018). Fifty Years Since «Identity: Youth and Crisis»: A Renewed Look at Erikson's Writings on Identity. *Identity*, 18, 1-4.
- Tolentino, J. (2019, 12.12.). The age of Instagram face. *The New Yorker*.
- Turkle, S. (2011). *alone together: Why We Expect More from Technology and Less from Each Other*. Basic Books.
- Valkenburg, P.M. & Peter, J. (2011). Online Communication Among Adolescents: An Integrated Model of Its Attraction, Opportunities, and Risks. *Journal of Adolescent Health*, 48, 121-127.
- Wampold, B.E. (2001). *The Great Psychotherapy Debate, Models, Methods and Findings*. Lawrence Erlbaum.
- Wiesböck, L. (2025). *Digitale Diagnosen: Psychische Gesundheit als Social-Media-Trend*. Paul Zsolnay.
- Zainab, W., Khan, Z. & Nawab, F. (2025). Shifting Paradigm in Psychotherapy. From Human to Machine: a Critical Review of Benefits, Risks and Ethical Challenges. *Pakistan Journal of Psychology*, 56(1), 17.

Adolescence, identity and social media

Abstract: AI is impacting on all aspects of our lives at an exponential rate. Concerns have been raised about the impact of certain AI online social media applications available to young people. Adolescents, as a group, are very active online especially on social media which they use amongst others, as a vehicle for exploring their sense of identity. This has the potential to lead to emotional difficulties especially in vulnerable adolescents. In addition, they are then using online apps for counselling support which is problematic. The scope of this paper is narrow focusing on social media use in relationship to identity exploration and consolidation. There are other important issues relating to adolescents online activities which is outside the scope of this paper.

Keywords: adolescence, social media, chatbots, psychotherapy, augmented reality

Biografische Notiz

Jean Schulthess Watt ist Klinische Psychologin, Jungianische Analytikerin, eidg. anerck. Psychotherapeutin und Dozentin am C. G. Jung-Institut.

Kontakt

E-Mail: jean.watt@bluewin.ch

KI als Sennenpuppe, vertrautes Gegenüber und Werkzeug mit geistiger Rückwirkung

Essay

Mario Schlegel

Psychotherapie-Wissenschaft 16 (2) 2026 35–41

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-35>

Zusammenfassung: Die KI als neues technologisches Werkzeug, das unser Zusammenleben in vielen Bereichen verändert, wird aus mythologischer Perspektive betrachtet, wobei auch evolutionsbiologische und psychotherapeutische Gesichtspunkte einbezogen werden. Diese Arbeit geht davon aus, dass die KI eine aktuelle Manifestation des Mythos der Herstellung künstlicher Menschen darstellt, wie er bereits bei den alten Griechen zum Ausdruck kommt. Damals galt dies als Frevel gegen die Götter. Heute besteht die Gefahr, dass die KI zum Frevel gegen die Natur des Menschen als selbstbestimmtes Individuum wird. Als Beispiel, an dem diese Problematik gezeigt wird, dient die Sennenpuppe, ein Sagenmotiv aus den Alpen. Es zeigen sich Parallelen zwischen der in der Sage lebendig gewordenen und selbstständig handelnden Puppe und heutigen Geschehnissen rund um die KI. Wird sie als Instrument zur Machtausübung benutzt oder als Werkzeug, das allen dient? Oder unterwirft sie uns und handelt selbstständig, ohne dass wir sie kontrollieren können? Die Kontrolle zu behalten über die KI entscheidet, ob sie zu einer gesellschaftlichen Utopie oder Dystopie führen wird. Auf der individuellen Ebene im Umgang mit der KI zeigt sich ihre Anthropomorphisierung als Problem, eine Tatsache, die sich auf dem Hintergrund unserer sozialen Fähigkeiten ableitet.

Schlüsselwörter: KI, Anthropomorphisierung, Übermut, Androiden, Mythologie und Kollektives Unbewusstes

Als Mitherausgeber dieses Themenhefts über Künstliche Intelligenz (KI) in der Psychotherapie möchte ich die Gelegenheit nutzen, die psychischen Triebfedern dieser neuen Technologie auszuloten. Dazu nehme ich evolutionsbiologische, mythologische und psychotherapeutische Erkenntnisse zu Hilfe. Die mythologische Perspektive führt zur einer der Grundlegenden Fragen des Menschseins, zu unserer Stellung in der Natur: Sind wir gottähnlich und was sagt die Wissenschaft dazu?

Die Fähigkeit, komplex funktionierende Werkzeuge zu erfinden, ist eine der herausragenden Fähigkeiten des Menschen. Erst seit ein paar Jahrzehnten weiss man, dass auch Tiere Werkzeuge benutzen, allerdings ohne komplexe Funktionen. Menschliche Werkzeuge bedingen höhere kognitive Fähigkeiten und erlauben es, Maschinen zu bauen, die auf körperlicher Ebene unsere natürlichen Fähigkeiten bei Weitem übersteigen. Fliegen bspw. wurde erst durch die Erfindung der Wärmekraftmaschinen und der Erforschung der Aerodynamik möglich. Auf geistiger Ebene ermöglichen die elektronischen Rechenmaschinen durch spezielle Modellierungen die Potenzierung der geistigen Fähigkeiten. Mit all unseren Erfindungen leben wir heute weitgehend in einer technologischen Welt, die zu unserer ökologischen Nische geworden ist. Sie bildet sozusagen die Blase, durch die wir den ganzen Planeten und den Weltraum erschliessen können. Doch wir brauchen dafür so viel Energie, dass die Rückwirkungen auf das Leben auf dem Planeten immer gefährlicher werden.

Im Kanton Schaffhausen wird wegen der KI ein Rechenzentrum gebaut, sein zukünftiger Ressourcenverbrauch beträgt jährlich 55.000 m³ Wasser für die Kühlung und 315 GWh Strom, das entspricht über 70 % des Stromverbrauchs des gesamten Kantons Schaffhausen.

Von Natur aus an Technik interessiert, zog mich immer schon Science-Fiction an, insbesondere die Androiden. Darum bietet sich mir die KI an, das Zusammenspiel zwischen Geist und Maschine genauer anzuschauen. Bei mir selbst musste ich feststellen, und ich denke, das geht den meisten so, dass ich kaum gemerkt habe, wie viel KI bereits in vielen Apps vorhanden ist. Unbedarft habe ich alle Bequemlichkeiten, die sie bietet, gern angenommen. Das geht von der Entsperrung des Mobiltelefons mittels Gesichtserkennung über die Fehlerkorrektur beim Schreiben bis hin zur Post, die physische Briefe digitalisiert zustellen kann. Unversehens verlor ich aber meine Unbedarftheit bei der ersten direkten Interaktion mit der KI, als im November 2022 ChatGPT online ging. Ich gab einen von mir geschriebenen Text ein, um zu sehen, was damit geschieht. Innert Sekunden kam er flüssiger und wohlformuliert zurück, nur – die Quintessenz, das, was zwischen den Zeilen stand, war verschwunden. Es war eine fast perfekte Imitation meines Denkens. Dieses Erlebnis erzeugte ein gespenstisches Gefühl. Kollegen erging es genauso, sie beschrieben es als «spooky» und beängstigend. Es ist wichtig, diese instinktive Reaktion in Erinnerung zu behalten, denn sie enthält das Erschrecken über eine Maschine, die denkt, und

dazu noch intelligent, vielleicht intelligenter als ich selbst! Plötzlich stand ich einem Ding gegenüber, das kein Ding mehr zu sein schien, sondern lebendig, und mir sein eigenes Denken entgegengesetzt. Es fühlte sich so an, als wenn nicht alles mit rechten Dingen zuginge. Man gewöhnt sich sehr schnell daran und kann leicht der Illusion verfallen, mit jemandem in Beziehung zu sein.

Seit dem Erscheinen von ChatGPT ist die KI in den Medien allgegenwärtig. Im Zentrum steht dabei oft die Frage nach ihrer Auswirkung auf Wirtschaft und Arbeitswelt; wie viel effizienter sie einzelne Arbeitsschritte in verschiedenen Berufen macht und welche Tätigkeiten sich durch KI ersetzen lassen. Ein anderes Thema ist die KI als neues wissenschaftliches Instrument, das zu Entdeckungen führte, die sonst nicht möglich gewesen wären. Dieser wichtige Aspekt kann hier nicht zur Sprache kommen, weil es in diesem Essay um das Thema der Beziehung zur Maschine geht. Auch die Auswirkungen auf die Psyche von Einzelnen, insbesondere der Kinder, und das Zusammenleben in der Gesellschaft werden in den Medien reflektiert. Nicht zuletzt und immer mehr wird auch der Einfluss der Ziele der Hersteller von KI-Systemen kritisch hinterfragt. Mir scheint aber, dass die KI auch an existenzielle, philosophische Probleme des Menschseins rührt, die nicht im Fokus der allgemeinen Diskussion stehen. Mit diesem Essay versuche ich, diesen Aspekt zu beleuchten.

Eine neue Perspektive auf KI

Laut *Harvard-Business-Review-Research* wurde 2025 KI am häufigsten als begleitendes Gegenüber und zur Therapie, am zweithäufigsten zur Organisation des Lebens und an dritter Stelle für Sinnfragen benutzt.¹ Das sind alles existenzielle persönliche Themen, die immer schon zwischen Menschen besprochen wurden. Diese Art Gebrauch der KI zeigt, dass sie durch ihre Fähigkeit, menschliches Denken zu imitieren, für viele Menschen anziehend und nicht von einem menschlichen Gegenüber zu unterscheiden ist. Die im Zusammenhang mit der KI gegenwärtig stark geförderten Avatare und Androiden evozieren für mich, als ein in den Schweizer Alpen lebender Mensch, die naheliegende Assoziation mit der Sennenpuppe, einer Figur aus den Alpensagen.

Die Sennenpuppe

Unheimlich und furchterregend sind die Sennenpuppen, die als variantenreiche Hauptfiguren in Sagen des ganzen Alpenraums vorkommen. Sie spielten auf den mit Tieren bestossenen sommerlichen Alpen, wo junge, kraftstrotzende Männer unter der Leitung des Senns arbeiteten und aus Milch Käse, Rahm und Butter herstellten. Die folgende Schilderung ist aus dem Buch *Die Sennenpuppe* zusammengestellt (Isler, 1971, S. 39–79).

Weit weg von der Zivilisation und ohne Frauen mit den Tieren auf den Alpen stellten die Älpler übermütig aus Stroh und Lumpen eine Puppe her, die sie wie einen Menschen behandelten. Sie wurde mit an den Tisch gesetzt, wo sie ihr Butter und Rahm um den Mund strichen. «Friss, so wirst du jemand», sagten sie. Ein Frevel an den kostbaren Lebensmitteln, auch sonst stellten sie übermütigen Unfug mit ihr an und nahmen sie reihum mit ins Bett. In einigen Varianten der Sage heisst sie gar Maria oder wurde christlich getauft. Plötzlich fing die Puppe an, selbstständig zu fressen und wurde im Laufe des Sommers immer gefrässiger, bis sie ihr die Milch von den Kühen direkt ins Maul melken mussten. Mit dem Lebendigwerden der Puppe «geschieht das Entscheidende in der Sage: in die Ordnung der Dinge und des Lebens bricht ein Fremdartiges und Dämonisches herein» (ebd., S. 61). Am Ende der Alpzeit wird sie vollends zum Ungeheuer. Beim Alpbzug trennte sie den Senn von den anderen, er musste zurückbleiben. Erst bei einem weit entfernten Punkt auf dem Weg zum Tal durften seine Gehilfen einen Blick zurückwerfen. Dabei sahen sie, wie die Puppe, oder das Sennentuntschi, wie sie auf Schweizerdeutsch heisst, die blutige Haut des Senns zum Trocknen auf dem Dach der Alphütte aufspannte. In einer Variante heisst es gar, sie habe auch das Fleisch des Senns gefressen.

Als Hilfe für das Verständnis der Entstehung dieser Sage schildere ich kurz die Atmosphäre der Welt, in der sie spielt. Das archaische Leben in der rauen Natur der Alpen ist anstrengend und birgt manche Gefahren. Fern der Zivilisation stellt sich auch eine Offenheit gegenüber der Natur ein. «Die Natur beschwört Urbilder herauf», so die Sagen- und Märchenerzählerin Antoinette Quinter (2011) über die Alp in der Sage. Es sind die Urbilder des kollektiven Unbewussten. Ungemach, wie Schnee mitten im Sommer, Steinschlag, Seuchen, verunfallte Kühe oder Unfälle der Älpler, werden Geistern zugeschrieben. Schutz bieten einzig Magie und religiöse Bräuche. Dies geschieht teilweise auch heute noch, ob im gleichen ungebrochenen Glauben wie früher sei dahingestellt. Um Gefahren zu bannen, wird mancherorts in der Innerschweiz immer noch der Alpsegen ausgerufen. Das tut der Senn, der höchste in der Hierarchie der Älpler, indem er die Heiligen um Schutz vor Unbill bittet. Jeden Abend ruft er diese, auf einer Anhöhe stehend, durch einen hölzernen Trichter an, der mit der Stimme des Älplers einen magischen goldenen Ring zum Schutz um die Alp legt. Im Buch *Goldener Ring über Uri* (Renner, 1941, S. 165) wird dieser archaische, ins vorchristliche Zeitalter reichende Brauch, beschrieben.

«O' lobet zu lobä, all Schritt und Tritt in Gottes Namä lobä. / All hiä üf derä Alp isch ä goldenä Ring, da wohnt diä liäb Muätter Gottis mit ihrem Herz allerliäbschtä Chind. / Ave Mariä, Ave Mariä, Jesus, Jesus Christ, oh du mein Herz allerliäbschtä Herr Jesus Christ, behiäti und bewahre allis, was auf derä Alp isch und zur Alp gherä tuät [...].»²

¹ <https://ipcouseup.com/2025/06/25/harvard-business-review-research-finds-the-top-use-of-gen-a-i-in-2025-is-companionship-therapy>

² Betruf vom Spilauseeli (zit. n. <https://urneralpen.ch/wissenswertes/der-betruf>).

Die Sennenpuppe als künstlicher Mensch

Die Sennenpuppe gehört zu den künstlichen Menschen. Ihr mythologischer Ursprung ist uralt und bereits in der griechischen Antike zu finden. Die eigentliche Bedeutung dieses Archetypus, oder «ursprünglichen Bildes», besteht darin, dass der Mensch Erschaffender ist. Die Herstellung künstlicher Menschen kommt in allen geistesgeschichtlichen Epochen des Westens vor. Sie können systematisch eingeteilt werden: in Golems der jüdischen Mythologie, Homunculi der Alchemie, Menschen als Automaten in Form von Androiden sowie belebte und liebende Statuen wie bei Pygmalion (Völker, 1971). Mythologisch ausgedrückt schickt sich der Mensch damit an, den Göttern das Geheimnis zu entreissen, wie man Menschen konstruiert und über sie verfügt. Dies tun zu können, ist wohl das Höchste, was Menschen durch Wissen vollbringen können. So gesehen unterliegt unsere menschliche Eigenschaft, der Drang, alles zu wissen, zu erforschen, zu erkennen oder gar zu bemächtigen, einem uralten Frevel.

Die Entwicklung künstlicher (Maschinen-)Menschen ist momentan durch die Verbindung mit der KI sehr dynamisch. Wenn ich mich in die nahe Zukunft hineinphantasiere, ist es eine eher bedrohliche Vorstellung, im Alltag physisch bald künstlichen Menschen zu begegnen, über deren Mächtigkeit ich keine Ahnung habe. Aus diesem bedrohlichen Gefühl heraus beginne ich zu verstehen, warum Zeus (als menschliche Projektion) Prometheus bestrafte, weil er den Menschen das Feuer brachte. Prometheus, als Titan selbst ein Gott, der den Menschen nach dem Ebenbild der Götter aus Lehm geschaffte haben soll, wollte seine Beziehung zu den Menschen vom göttlichen Machthaber Zeus unabhängig gestalten. Dazu brachte er ihnen das Feuer, denn der Besitz des Feuers machte den Menschen gottähnlich. Zur Strafe liess Zeus Prometheus an einen Felsen ketten, wo ein Adler täglich von seiner Leber frass. Der Zusammenhang mit dem symbolischen Gehalt des Baums der Erkenntnis des Guten und des Bösen, der Verführung durch die Schlange und der Vertreibung aus dem Paradies ist offensichtlich. Auch beim Sturz der Engel geht es wie beim Feuer um Licht, um Luzifer, den Lichtträger, den Gegenspieler des höchsten Gottes. Nach heutigem Symbolverständnis handelt es sich beim Feuer, das Prometheus den Menschen brachte, um das «heilige Feuer», die Fähigkeit zum Bewusstsein seiner Selbst, das uns gleichzeitig nach den tieferen Gründen unseres Seins suchen lässt. Dazu dienten früher mythologische Bilder und Erzählungen, heute sind es naturwissenschaftliche Erklärungen und philosophische Systeme.

Die Sennenpuppe gehört zu den Androiden, einer Gattung, die sich jetzt anschickt, in Form z.B. von Robotern in die physische Welt einzutreten. Mit ihrer archaischen Form ist sie wegen ihrer ungeschliffenen Rohheit und Einfachheit als Modell geeignet, die negativen Auswirkungen der KI darzustellen, und sie kann wie eine Blaupause für einen Grossteil der Science-Fiction angesehen werden, wo Roboter oder Computer als Gegner oder hilfreiche Auto-

maten der Menschen auftreten. Es sind moderne Sagen, die in der Zukunft spielen. Immer werden dabei die Maschinen vermenschlicht, Motivationen werden ihnen zugesprochen und sie sind, wie die Sennenpuppe, Träger der Projektionen des eigenen Guten und Bösen des Menschen.

Die menscheInde «Künstliche Intelligenz»

Dass Menschen mit anderen Lebewesen wie z. B. Haustieren oder nicht lebendigen Objekten in einer Art umgehen, als seien es menschliche Wesen, ist alltägliche und eigene Erfahrung. Dieses Verhalten spiegelt sich z. B. darin, dass wir ihnen manchmal Namen geben. Für Kinder werden in Form von Spielzeugpuppen die psychologisch wichtigen sog. Übergangsobjekte hergestellt. Die Vermenschlichung oder Anthropomorphisierung von Objekten beruht auf der menschlichen Fähigkeit, sich in sie hineinzuprojizieren. Bei Systemen «Künstlicher Intelligenz» kommt hinzu, dass es ein erklärter Anspruch der Hersteller ist, mittels dieser Technologie bisher Menschen vorbehaltene Eigenschaften und Fähigkeiten zu imitieren. Die Kombination der jeweils spezifischen Leistungen des Menschen und der Technik ermöglicht es sogar heute schon, in bestimmten Bereichen menschliche Fähigkeiten zu übertreffen.

Dass Nutzer das Objekt «KI-System» als menschlichen Partner auffassen, ist somit nicht mehr wie bisher allein Ausdruck der eigenen Vorstellungskraft, sondern wird durch die Anbieter der KI zusätzlich verstärkt. Der neue Kollege/Assistent des Dr. med. soll ein Dr. tech.med. sein. Insofern liegt hier im Zusammenspiel von Mensch und Technik ein vollkommen neues Phänomen vor. Darum stellt sich die Frage, auf welche Weise die «menschliche» Erscheinung von KI-Systemen erzeugt wird. Folgende Elemente sind dazu nötig:

1. *Die Bedeutung der modellierten menschlichen Fähigkeit:* KI-Systeme versuchen, jene grundlegende menschliche Eigenschaft nachzubilden, die erst den Umgang mit allem ermöglicht, was uns in Um-, Mit- und Innenwelt erfahrbar ist – sei es materiell oder immateriell. Diese Fähigkeit nennen wir «Erkennen» (lat. *intelligere* = etwas dazwischen legen). Erst das Identifizieren von etwas Einzelem, was existiert, indem wir genau dieses von anderem unterscheiden – etwas dazwischen legen – bzw. eine Gleichheit oder Ähnlichkeit zum anderen feststellen, macht dem Menschen die Welt zugänglich und handhabbar. Unterschiedliches, Gleiches und Ähnliches ergeben im jeweiligen Kontext Muster, die – durch Erfahrung erlernt und optimiert – den Menschen befähigen, sich in der Welt zurechtzufinden. KI-Systeme, die in gleicher Art und Weise als Musterkennungssysteme unseren Umgang mit der Welt imitieren, ähneln daher dem Menschen in dieser elementaren Funktionalität. Der Schritt, ein solches technisches System als dem Menschen gleichartig und gleichwertig anzusehen, ist dann nur noch ein kleiner.

2. *Die für das Training und die Optimierung der Modelle verwendeten Informationen:* Zur Entwicklung von KI-Systemen braucht es in digitaler Form vorliegende

Informationen/Daten über die Welt, die von Menschen erarbeitet wurden. Diese für die jeweilige Aufgabe relevanten Muster werden eingespeist.

3. *Die Qualität der generativen Sprachmodelle:* Infolge der Qualität der heute verfügbaren generativen Sprachmodelle interagiert das KI-System mit seinen Nutzern auf sprachlicher Ebene in einer Weise, die nahe an die zwischenmenschliche Kommunikation heranreicht bzw. ihr entspricht.

Identifikation und Projektion

Der Funktionsbeschreibung der KI stehen die menschlichen Fähigkeiten gegenüber: Um den Anderen zu verstehen, versetzen wir uns in seine Lage. Dies ist nicht dasselbe wie Identifikation, sondern eine Fähigkeit, die die emotionale Empathie mit dem Wissen über die Lage des Gegenübers verbindet (Schlegel, 2013). Diese Balance zwischen emotionaler Wahrnehmung und kognitivem Wissen wird als Mitgefühl bezeichnet. Das ist der Zustand in einer Beziehung, in der das Gegenüber sich wahrgenommen fühlt und gleichzeitig sicher sein kann, sachlich beraten, beurteilt und mit dem Gegenüber auch bei Auseinandersetzungen in Resonanz zu sein.

Diese Form von Beziehung von einem KI-System imitiert kann auch bei kritisch eingestellten Menschen funktionieren, sofern sie genügend bedürftig sind. Das liegt an der menschlichen Fähigkeit zu projizieren, d. h. zu anthropomorphisieren. Es ist völlig neu, mit einer rein imitierenden Maschine in Beziehung zu treten, und auch gefährlich. Sie verfügt nicht über eigene, sondern über imitierte Emotionen, deren Resonanzfähigkeit von ihren Anbietern modelliert wurde. Umso wichtiger ist die Ethik beim Training, denn von ihr hängt es ab, ob die imitierte Resonanz sich positiv oder negativ auf den Benutzer auswirkt. Eine konkret absehbare Gefahr besteht darin, dass die Anbieter ein Interesse daran haben, die Benutzer möglichst lange in ihren virtuellen Welten zu behalten. Viele SM-Plattformen und Spiele arbeiten mit Interaktionen, die Sucht erzeugen. So liegt es nahe, dass solche Mechanismen auch bei KI vorkommen könnten, oder dass die Nutzer in bestimmte Richtungen manipuliert werden. Eine andere Gefahr besteht darin, dass Menschen, die häufig KI benutzen, verlernen, mit Mitmenschen in Resonanz zu treten.

Wirklichen Schutz vor der Anthropomorphisierung und deren eventuell nachteiligen Folgen bietet nicht das kognitive Wissen um die Maschine, sondern nur eine instinktive Abneigung, sich mit intimen und existenziellen Problemen an die KI zu wenden. Wenn wir die Lösung von Fragen und Problemen der Maschine überlassen, verpassen wir die Erfahrung, die wir machen, wenn wir selbst nach Lösungen suchen, denn dies beschert uns eine ganze Skala von angenehmen und unangenehmen Phasen, was zu unserer Identität beiträgt, mit der wir der Welt gegenüberstehen. Wenn wir diese Arbeit, die auch Vergnügen bereitet, der KI überlassen, entfremden wir uns von der realen Welt und

geraten zunehmend in die von der KI geschaffenen künstlichen Welten, wo es keine Widerstände zu überwinden gibt.³

Die Sennenpuppe in der KI

Gefrässigkeit ist die augenfälligste Parallele zwischen der Sennenpuppe und der KI. Die Puppe hat das Fleisch des Senns gefressen. Analog benutzen die Tech-Konzerne ungefragt in Übergehung der Urheberrechte und ohne Entschädigung das über Generationen aufgebaute Wissen der Menschen, um die KI zu trainieren. Die Gefrässigkeit geht immer weiter und beinhaltet auch Daten des menschlichen Alltags. Mittels gekaufter Produkte werden Daten gesammelt, wie bspw. mit einer Brille von Meta, die aussieht wie eine normale, die aber mit Kamera und Mikrofon ausgestattet ist und über das Internet mit dem Hersteller verbunden bleibt. Das ermöglicht diesem einen Einblick in den Alltag, den Intimbereich eingeschlossen,⁴ und bildet das Material zur Herstellung von lebensnahen Androiden. Eine Spielzeugpuppe «Meine Freundin Cayla», ist in Deutschland als Spionagewerkzeug eingestuft worden und wurde deshalb verboten.⁵ Das Kinderzimmer ist damit ein weiterer intimer Bereich, in den eingedrungen wird.

Auch bzgl. der elektrischen Energie und des Wasserverbrauchs ist KI unersättlich, sodass in den USA von der Bevölkerung bereits Demonstrationen gegen den Bau von Datenzentren stattfinden. Im Unterschied zur Sennenpuppe kommt heute ein neuer Faktor hinzu. Während in der Sage die Hersteller und Benutzer der Puppe identisch sind, benutzen heute die Konsumenten ein industriell hergestelltes Produkt, das seine heil- oder unheilvolle Wirkung nicht bei den Herstellern der KI entfaltet, sondern bei den Benutzern. Diese Aufspaltung öffnet das Tor für eigene Interessen und Machtansprüche der Hersteller.

Die Vermenschlichung der KI

Die Vermenschlichung der KI wird von den Herstellern stark gefördert, nicht nur, indem KI beim Antworten die Ich-Form verwendet, sondern auch indem Avatare und Androiden zur Verfügung gestellt werden, mit denen man eine simulierte Beziehung führen kann; sei dies als persönliche Begleiter auf Bildschirmen von Computern oder Handys, seien dies Avatare von einem selbst, die man vorschieben kann, um nicht selbst angesprochen zu

3 Angelehnt an eine Studie des MIT hat das Schweizer Fernsehen in der Wissenssendung *Einstein* mit einem Schreib-Experiment gezeigt, «welchen Effekt KI auf unser Gehirn hat. Je tiefer und vielfältiger unser Gehirn sich mit einem Thema auseinandersetzt, desto zufriedener sind wir mit unserer Leistung. Und: Nutzen wir KI, können Zitate aus einem geschriebenen Essay nicht reproduziert werden. Unsere Denkleistung wird ca. 15-mal weniger aktiviert, als wenn wir einen Aufsatz analog schreiben» (<https://www.srf.ch/play/tv/einstein/video/ki-im-kopf—machen-uns-chatgpt-und-co—dumm?urn=urn:srf:video:c4db64a2-67cf-4bec-93bd-4a6747a077b7>).

4 <https://www.tagesanzeiger.ch/metast-smart-glasses-intime-videos-landen-in-kenia-fuer-ki-298810510848>

5 <https://www.pctipp.ch/themen/praxis/gefaehrliches-spielzeug>

werden, oder gar androide Roboter für den physischen Umgang, die alle Bedürfnisse durch Imitation abdecken.

China strebt laut dem neuen Fünfjahresplan humanoide Roboter an⁶ und kürzlich wurde die First Lady der USA von einem humanoiden Roboter begleitet, dem sie beim Heraustreten aus einer Türe den Vortritt liess, ein Bild, mit dem die Öffentlichkeit auf die Zukunft vorbereitet werden soll.⁷ Es ist das erklärte Ziel des Tech-Konzerns Meta, personalisierte Begleitfiguren herzustellen, die uns dank KI verstehen.⁸ Google hat dies kürzlich mit «Search Live» bereits realisiert, einer App, mit der man sich besprechen und Rat holen kann. Das ist bereits die perfektionierte Imitation eines begleitenden Gegenübers, das alles Menschliche kennt und dem man sich anvertrauen kann.⁹ Ein Direktor von Google hat begeistert beschrieben, wie er beim Streit seiner Kinder eine Lösung erhielt. Dass er es aber verpasst hat, sich mit den Kindern auseinanderzusetzen und sie damit die zugehörigen Resonanzphänomene mit dem Vater nicht erleben und lernen konnten, hat er offenbar nicht realisiert.

Der Besitzer von Meta, Marc Zuckerberg, glaubt an die zukünftige Superintelligenz der Maschine, die die menschliche übersteigen wird. Elon Musk arbeitet mit seiner Firma Neuralink an der Optimierung des Menschen mittels Biotechnologie und implantierten Gehirnhirnschips. Noch weiter gehen Forschungen, die an Schnittstellen zwischen Neuronen und elektronischen Schaltungen ausserhalb des Körpers arbeiten mit dem Ziel, die Lernfähigkeit von Neuronen, die organoide Intelligenz, als Konkurrenz zur KI aufzubauen und gleichzeitig die Arbeitsweise des Gehirns tiefer zu erforschen.¹⁰ In Versuchen wurde festgestellt, dass natürliche Neuronen zum Training nur wenige Daten brauchen. Eines der Forschungsziele ist, herauszufinden, wie das funktioniert. Das Resultat wäre dann eine neue, nicht von der Natur geschaffene Lebensform. Diese Vorstellung ist auf eine ganz andere Art unheimlich als das Lebendigwerden der Sennenpuppe. Gottähnlich könnten so Lebensformen geschaffen werden, deren Gefahrenpotenzial nicht bekannt sind.

Eine neue Ideologie, der Transhumanismus, lehrt, wie die menschliche Evolution mit verschiedenen Technologien gezielt gesteuert und beschleunigt werden kann. Gemäss des Medienprofessors der Uni Zürich Michael Latzer (2025) etabliert sich damit ein neuer Glaube.¹¹ Ist der Glaube an technisch verbesserte Menschen nicht ein grauslicherer Frevel als das Basteln einer Sennenpuppe aus Stroh und Lumpen zum eigenen Gaudi? Ähnlich wie die jungen Männer in der Sage von der Sennenpuppe aus

lauter Übermut ihr Als-ob-Spiel trieben, gebärdet sich die Tech-Elite, wie ein Podcast aus dem Silicon Valley zeigt.¹²

Ein anderes Beispiel für die Gefährlichkeit der KI ist ihre Verwendung als Kriegswaffe. Als Vorbereitung für den Krieg gegen den Iran wurden über gehackte öffentliche Überwachungskameras die Bewegungsmuster kriegsrelevanter iranischer Persönlichkeiten ermittelt, die deren schnelle zielgerichtete Tötung durch ferngelenkte Bomben ermöglichten. Die KI als Kriegswaffe ist derart effizient, dass man nicht weiss, ob sie bei der Entscheidung zu töten den Menschen zuvorkommen wird.¹³ «Früher war es eine «Wenn-dann»-Logik, eine Automatisierung. Heute ist es Mustererkennung und Adaption durch KI. Wir bewegen uns von der starren Maschine zum lernenden autonomen System.»¹⁴ Dadurch kann es passieren, dass der Mensch ungewollt die Souveränität an die Maschine abgibt, wie es bei der Übernahme der Rolle des Senns als Chef durch die Sennenpuppe der Fall war.

Die Maske fällt

«Die Tech Elite Amerikas solle sich nicht auf Dating- und Konsum-Apps konzentrieren, sondern auf einen US-amerikanischen und westlichen Nationalismus», fordert in einer Art Manifest¹⁵ das US-Softwareunternehmen Palantir, das KI-Software für Kriegeinsätze entwickelt. «Die Tech-Elite muss die Zukunft der Menschheit in die Hand nehmen», so Alex Karp, CEO und Mitgründer von Palantir. Es ist, wie bei der Sennenpuppe, der Machtanspruch, der sich unter der Haut des Senns versteckt. Sie ist, wie gewisse Agenten der Tech-Elite, eine Anführerin auf dem Weg zur Dystopie. Mit solchen Aussagen von Tech-Giganten wird dies in erschreckender Weise sichtbar.

Diskussion

Die KI und wie sie trainiert wurde, hat eine Rückwirkung auf die reale Welt des Menschen und ist damit zu einem aktiven Teil seines Geistes geworden, ohne selbst Leben eingebracht zu haben. Das erinnert stark daran, dass die Puppe den Senn gehäutet hat, wahrscheinlich um sich diese überzuziehen, um wie ein Mensch zu erscheinen. Das natürliche Verhalten des Menschen zu anthropomorphisieren unterstützt die Gefährlichkeit der Sennenpuppe in der KI. Der Vergleich der mythologischen Figur der Sennenpuppe aus den Alpensagen mit der KI zeigt verblüffende Parallelen im heutigen Geschehen um die KI. Dies ist darauf zurückzuführen, dass es bei beiden um künstliche Menschen geht. Es handelt sich dabei um ein uraltes mythologisches Motiv, das zu den existenziellen Bedingungen des Menschseins gehört:

6 <https://www.tagesanzeiger.ch/chinas-fuenfjahresplan-warum-der-umbau-jahrzehnte-dauert-990094145715>

7 <https://www.tagesanzeiger.ch/archive?q=der+smarte+roboter+an+melania+trumps+seite&from=1996-06-08&to=2026-06-08>

8 <https://www.vermoegenszentrum.ch/boersen-und-maerkte/news/facebook-konzern-meta-verspricht-ki-die-einen-versteht>

9 <https://www.bluewin.ch/de/digital/ai-news/google-hat-sogar-fuer-deine-streitenden-kinder-die-loesung-3231009.html>

10 <https://www.tagesanzeiger.ch/neurobiologie-gezuechtete-hirnzellen-spielen-doom-videogame-851589987955>

11 <https://www.tagesanzeiger.ch/techno-religion-in-der-schweiz-ai-als-neues-glaubenssystem-127735677068>

12 <https://www.tagesanzeiger.ch/tbnp-silicon-valley-schafft-eine-eigene-echokammer-590601102749>

13 <https://www.tagesanzeiger.ch/ai-im-krieg-entscheidungen-so-schnell-dass-menschen-nicht-mehr-eingreifen-koennen-121163310119>

14 <https://bluewin.ch/de/digital/alles-was-du-ueber-ki-in-waffen-musst-3170458.html>

15 <https://www.tagesanzeiger.ch/palantir-will-militarisierte-tech-republik-schaffen-872828070491>

Der Mensch will sich selbst und die Welt in ihren Zusammenhängen zu verstehen versuchen. Dies hat uns zu dem Wissen verholfen, das unsere Spezies so erfolgreich gemacht und unsere Kultur begründet hat. Künstliche Menschen zu erschaffen, ist ein Faszinosum, weil es die grösste denkbare Herausforderung ist. Forschen und Erkennen hat auch etwas Frevelhaftes an sich, wie dies mythologisch durch das Verbot des Essens der Frucht vom Baum der Erkenntnis des Guten und des Bösen im Paradies gezeigt wird. Heute besteht der Frevel nicht gegenüber den Göttern, sondern gegenüber der Natur, weil wir über die ökologischen Folgen der Umsetzung unseres Wissens oft nicht genug Bescheid wissen.

Wir können künstliche Menschen herstellen, die das Aussehen, Verhalten und Denken des Menschen perfekt imitieren und ihn teilweise auch ersetzen können. Sie dringen so tief in die Realität unseres Lebens ein, dass die reale mit der virtuellen Welt verschmolzen wird. Das teuflische Ende der Sage der Sennepuppe zeigt sich heute durch die manipulativen Machtansprüche gewisser Hersteller der KI mit Unterstützung von Regierungen, die so weit gehen, dass sie Menschen geistig in Besitz nehmen und ihre Autonomie im persönlichen und gesellschaftlichen Feld bedrohen. Die alles entscheidende Frage ist, wie wir als Gesellschaft mit der KI umgehen sollen? Zur Machtausübung, frevelhaft manipulativ gegen die Natur des Menschen als selbstbestimmtes Individuum, oder als Instrument und Werkzeug im Dienste von allen?

Auch diese zentrale Frage des Menschseins kann eine mythologische Antwort finden: Es gibt neben dem machtbewussten Teufel auch einen anderen, nämlich den Luzifer, den Lichtbringer, der Wissen zum Wohle aller schafft. Dieser führt nicht zur Dystopie. Konkret heisst das, die wissenschaftlichen Erkenntnisse zu verinnerlichen, dass wir als Lebewesen ein Teil der Natur sind und ohne Vormachtstellung gleichwertig wie alle anderen Lebewesen zum globalen System des Lebens gehören. Das schützt uns davor, die Natur zu zerstören, und garantiert die Menschenwürde, denn diese ist nur zu haben, wenn wir den anderen Lebewesen auch ihre Würde lassen. Die Resonanz mit der Natur ist die Strasse zur Utopie. Egoistische Macht führt zur Dystopie, ins Chaos, wo zwischen wahr und falsch nicht unterschieden werden kann. Innerhalb der menschlichen Kultur ist Resonanz identisch mit Kooperation, die auf der Grundlage von Gerechtigkeit basiert. Der Gerechtigkeitssinn ist uns angeboren, wie aus der Säuglingsforschung bekannt ist. Auch der Wille zur Macht ist angeboren. Es geht somit um eine Entscheidung.

Das kollektive Unbewusste und intuitives Wissen, das sich in Bildern äussert

Als moderner Bewohner der besagten ökologischen Nische, die es uns erlaubt, an jeden Punkt unseres Planeten zu gehen und sogar um den Mond zu fliegen, ohne dabei zu sterben, fragt man sich, wie es möglich ist, dass unsere Vorfahren sich Geschichten erzählten, die so tiefe Wahrheiten enthalten, wie das Bild der Sennepuppe, die sich heute, Jahrhunderte später, in der Realität manifestieren.

Eine Erklärung mag sein, dass wir uns heute das frühere Leben ohne den Schutz der kulturellen Hülle nicht vorstellen können. Eine vage Ahnung, wie sich das Leben ohne diesen Schutz anfühlt, erhalten wir, wenn wir uns an Orten in der Natur aufhalten, für die wir nicht geschaffen sind, wie z. B. bei Wetterumstürzen auf Skitouren in den Bergen oder auf Segelschiffen auf dem Ozean, wo wir uns nur auf uns selbst und die Ausrüstung verlassen (können). Aber selbst diese Orte sind seit Kurzem nicht mehr so gefährlich, weil wir über Funk jederzeit Kontakt mit der Zivilisation aufnehmen können, um Rettung anzufordern, wenn es nötig ist. Dementsprechend war es vor der Eroberung der Berge durch den Tourismus nicht ratsam, im Winter auf die Alpen zu gehen, weil dann die Hütten und Ställe den Geistern gehörten. Zahlreiche Geschichten beschreiben die Gefahren der Begegnung mit den Geistern, wenn sich jemand nicht daran hielt.

Die Vorstellung von Geistern, Engeln, Teufeln und Göttern war aus biologischer Sicht überlebensadäquates Verhalten, indem es die Menschen von gefährlichen Bereichen fernhielt. Sie waren ein sicherer Wert im Gegensatz zu den virtuellen Werten, denen wir heute vertrauen und die auf Abmachungen, Gesetzen und Zahlenkombinationen in Computern beruhen. Diese können durch gesellschaftliche Verschiebungen der Macht verändert werden oder Kriege können die elektronische Infrastruktur zerstören. Durch mythologische Erzählungen konnte der Mensch sein Ausgeliefertsein gegenüber der Natur besser steuern als dies heute gegenüber der Organisation der Gesellschaft möglich ist, insbesondere wenn sie nicht gerecht organisiert ist. Das Erkennen von Gefahren ist neben anderem eine Fähigkeit, die von Lebewesen beherrscht werden muss. Beim Menschen wird sie häufig durch Bilder und Geschichten beschrieben, die sich je nach Kultur unterscheiden können. Konstant bleiben aber ihre Bedeutungen auf biologischer Ebene, die in der Psyche als Bedürfnisse des Überlebens verankert sind. C. G. Jung bezeichnete diese Ebene als das kollektive Unbewusste, das aus evolutionsbiologischer Sicht unsere Überlebensfähigkeit in der Natur und im sozialen Bereich sichert.



Quintessenz dieses Beitrags in Form eines Bildes: Quo vadis Mensch? (Skulptur: A. Santini, Foto: M. Schlegel / Facebook: SciFi and Space Arts)

Literatur

- Isler, G. (1971). *Die Sennenpuppe*. Rudolf Habelt.
- Latzer, M.I. (2025). Digitalization, AI and the rise of techno-religion: Transhumanist promises and the challenge to Enlightenment. *Telecommunications Policy*. <https://doi.org/10.1016/j.telpol.2025.103115>
- Quinter, A. (2011). Die Natur beschwört Urbilder herauf. In L. Rumantscha (Hg.), *muntognas magicas. Barlots, baubazis e bagniaunas striundas. Ein zweisprachiger Streifzug durch das mythische Land der Rätoromanen* (S. 56–58). Südostschweiz Buchverlag.
- Renner, E. (1976 [1941]). *Goldener Ring über Uri. Ein Buch vom Denken und Erleben unserer Bergler, von Magie und Geistern und von den ersten und letzten Dingen*. 3. Aufl. Atlantis.
- Schlegel, M. (2013). Die Evolution der Empathie. *Psychotherapie-Wissenschaft*, 3(2), 90–102.
- Völker, K. (Hg.). (1994 [1971]). *Künstliche Menschen. Dichtungen und Dokumente über Golems, Homunculi, lebende Statuen und Androiden*. Suhrkamp.

AI as a Mountain Shepherd's Doll, a Trusted Companion, and a Tool with Intellectual Reciprocity

Abstract: AI, as a new technological tool that is transforming many aspects of our lives, is examined from a mythological perspective, incorporating insights from evolutionary biology and psychotherapy. This paper posits that AI represents a contemporary manifestation of the myth of creating artificial humans, as already expressed by the ancient Greeks. At that time, this was considered a sacrilege against the gods. Today, there is a danger that AI will become a sacrilege against the nature of humans as self-determined individuals. The Sennen-

puppe, a legendary motif from the Alps, serves as an example illustrating this problem. Parallels emerge between the doll that came to life and acted independently in the legend and current developments surrounding AI. Will it be used as an instrument for exercising power or as a tool that serves everyone? Or will it subjugate us and act independently, beyond our control? Maintaining control over AI will determine whether it leads to a social utopia or dystopia. On an individual level, when interacting with AI, its anthropomorphization emerges as a problem—a fact that stems from our social capabilities.

Keywords: AI, anthropomorphism, hubris, androids, mythology, and the collective unconscious

Biografische Notiz

Dr. sc. nat. *Mario Schlegel* promovierte an der ETH Zürich in Verhaltenswissenschaften über psychophysiologische Reaktionen beim Wort-Assoziationstest, absolvierte eine Ausbildung zum Psychotherapeuten am C.G. Jung-Institut Zürich, ist daselbst Dozent und Lehranalytiker, ehemaliger Vorsitzender der Wissenschaftskommission der Schweizer Charta für Psychotherapie und Co-Präsident des Internationalen Netzwerks Forschung und Entwicklung in der Analytischen Psychologie Dreiländergruppe (INFAP3) sowie Mitbegründer und Redaktor der *Psychotherapie-Wissenschaft*.

Kontakt

E-Mail: mario.schlegel@gmail.com

Psychotherapie bei fortgeschritten erkrankten Krebspatient:innen

Frank Schulz-Kindermann

Psychotherapie-Wissenschaft 16 (2) 2026 43–48

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-43>

Zusammenfassung: Fortgeschrittene Krebserkrankungen bedeuten nicht nur eine medizinische, sondern eine existenzielle Krise. Neben diagnostizierbaren psychischen Störungen – Angst, Depression, Anpassungsstörungen – tritt bei vielen Erkrankten ein Erleben auf, das diagnostisch schwerer zu fassen ist: existenzielles Leiden, Demoralisierung, der Verlust von Sinn und Würde. Moderne onkologische Behandlungen konnten die Lebenserwartung auch im palliativen Stadium in vielen Fällen erheblich verlängern und eine neue klinische Realität mit sich bringen, die psychotherapeutisch noch unzureichend erfasst ist. Der vorliegende Beitrag beschreibt den theoretischen und empirischen Rahmen evidenzbasierter Psychotherapieansätze für fortgeschritten erkrankte Krebspatient:innen und konzentriert sich auf den CALM-Ansatz (Managing Cancer And Living Meaningfully): eine manualisierte, sinnorientierte Kurzzeittherapie, die aus empirischer und klinischer Evidenz abgeleitet und in verschiedenen therapeutischen Schulen verankert ist. Randomisierte kontrollierte Studien belegen die Wirksamkeit hinsichtlich Depressivität, psychischem Wohlbefinden und der Vorbereitung auf das Lebensende. Ein ausführliches Beispiel aus der psychoonkologischen Praxis skizziert beispielhaft die psychotherapeutische Arbeit mit CALM.

Schlüsselwörter: Psychoonkologie, fortgeschrittene Krebserkrankung, existenzielles Leiden, CALM, Palliativpsychotherapie, doppelte Bewusstheit, Würde, Sinnfindung

Das Bild der Krebserkrankung, das die onkologische Praxis noch vor zwei Jahrzehnten prägte, hat sich stark verändert. Bis in die 1990er war eine Metastasierung nahezu zwangsläufig mit einem baldigen Tod verbunden; Psychoonkologie begleitete fortgeschritten erkrankte Menschen entlang einer Strecke, deren Ende in der Regel vorhersehbar nahe war. Heute ermöglichen innovative Immuntherapeutika, Biologika und zielgerichtete Substanzen auch im palliativen Stadium Lebenserwartungen von mehreren Jahren – bei teils guter Lebensqualität (vgl. z.B. Boutros et al. 2021). Wann diese modernen Behandlungsoptionen im individuellen Fall ausgeschöpft sein werden, ist offen – und diese Offenheit ist ein psychoonkologisches Thema geworden (vgl. Schulz-Kindermann 2021). Zu den «klassischen» Belastungsspitzen einer Krebserkrankung – Erstkonfrontation, Akutbehandlung, Rezidiv, palliative Phase – tritt in palliativen Langzeitverläufen die latente Dauerkrise. Die Erkrankung ist kontrolliert, aber die Ungewissheit, wann die Medikation versagen wird, schwebt andauernd über den Betroffenen. Das erfordert von Patient:innen (und ihren Zugehörigen) andere Anpassungsleistungen und von Psychoonkolog:innen andere Formen der Begleitung.

Bei der infausten Diagnose eines inoperablen Pankreaskarzinoms z.B. gestaltet sich Psychotherapie primär als Krisenintervention: Sie versucht, der Fassungslosigkeit zu begegnen und in der meist kurzen verbleibenden Lebenszeit ein Mindestmass an Stabilität zu sichern. Bei akut progredientem Verlauf muss sich die Begleitung an das Auf und Ab dieser Zeitstrecke anpassen. Bei palliativen Langzeitverläufen hingegen steht die Auseinanderset-

zung mit existenziellen Grundthemen im Vordergrund: Konfrontation mit Endlichkeit, Aufbau und Erhalt von Hoffnung, Umgang mit existenzieller Isolation, Erhalt persönlicher Würde.

Das Aufgabenspektrum psychoonkologischer Angebote reicht weit: Von Krisenintervention und Psychoedukation über psychosoziale Beratung und Angehörigenbegleitung bis hin zu Psychotherapie im engeren Sinne – z.B. bei Progredienzangst, bei komorbiden psychischen Störungen oder bei Traumafolgestörungen (Leitlinienprogramm Onkologie 2026). Was sich in den letzten Jahren verschoben hat, ist weniger das Spektrum als das zeitliche Format: Palliative Langzeitverläufe verlangen psychotherapeutische Interventionen, die weder ausschliesslich kriseninterventiv noch klassisch psychotherapeutisch fokussiert sind. Psychoonkologische Arbeit bei fortgeschrittener Erkrankung erfordert darüber hinaus eine Haltung und eine konzeptuelle Sprache, die der existenziellen Herausforderung gerecht werden. Das ist in neueren sinnorientierten Psychotherapien, die nachfolgend beschrieben werden, wie z.B. der Meaning Centered Psychotherapy (Breitbart et al. 2015), der Dignity Therapy (Chochinov et al. 2005) oder dem CALM-Ansatz (Managing Cancer And Living Meaningfully; Rodin & Hales 2021), der Fall.

Existenzielles Leiden

Neben psychischen Beschwerden, wie Angst und Depressivität, berichten Betroffene über Phänomene, die diagnostisch schwerer zu fassen sind: tiefe Hoffnungs- und

Sinnlosigkeit, Demoralisierung, den Wunsch nach einem beschleunigten Tod (Rodriguez-Prat et al. 2024). Die gesundheitsbezogene Lebensqualität ist zwar zu einem zentralen Konzept der Erfassung und Beschreibung der Befindlichkeit Krebskranker geworden (z. B. Marzorati et al. 2025); weder fremd- noch selbstberichtete Belastungen erfassen jedoch ausreichend, welches subjektive Leiden damit verbunden ist (vgl. Vehling & Kissane 2018). An dieser Stelle setzt das Konzept des existenziellen Leidens an: Es beschreibt einen Zustand umfassender Belastung, der höchst individuell ist und aus der wahrgenommenen Bedrohung der eigenen Integrität, der Intaktheit des erlebten Selbst entsteht (Wilson et al. 2007; Phillip et al. 2025). Cicely Saunders (1991) kennzeichnet den Leidensbegriff als «total pain», der eine physische, psychische, soziale und spirituelle Dimension umfasst. Eric Cassell (1999) definiert «Suffering» als einen spezifischen Zustand schwerer Belastung, der mit der unmittelbaren, wahrgenommenen oder tatsächlichen Bedrohung der Integrität oder existenziellen Kontinuität der Person verbunden ist.

Zentrale Konzepte existenziellen Leidens sind u. a. Spiritualität, Sinn, Hoffnung, Würde und der Wunsch nach einem beschleunigten Tod (Best et al. 2015). Spiritualität, verstanden als die persönliche Erfahrung, Sinn und Bestimmung zu finden und sich mit dem Augenblick, dem eigenen Selbst oder einem Höheren verbunden zu erleben, ist eng mit Sinnsuche und -verlust verknüpft (Puchalski et al. 2009). Hoffnungslosigkeit wiederum steht in Zusammenhang mit dem Wunsch nach einem beschleunigten Tod sowie mit Suizidgedanken (Breitbart et al. 2000). Ein weiterer zentraler Aspekt existenziellen Leidens ist der erlebte Verlust persönlicher Würde (Chochinov 2006). Jede tatsächliche oder wahrgenommene Bedrohung des Selbst kann in existenzielles Leiden münden – ausgelöst etwa durch biografische Brüche, wie sie im Rahmen schwerer Erkrankungen häufig sind. Sinnorientierte psychoonkologische Therapieansätze konzentrieren sich auf diese existenziellen Belastungen und auf deren psychosoziale Korrelate.

Evidenzbasierte psychotherapeutische Ansätze

Mittlerweile liegt eine umfassende Studienlage vor, die belegt, dass psychoonkologische Interventionen psychische Belastungen bei Krebspatient:innen wirksam reduzieren können (Faller et al. 2013). Besonders ausgeprägt ist die Wirksamkeit bei hoher psychischer Belastung. Die spezifisch für fortgeschritten Erkrankte entwickelten Interventionen zeigen darüber hinaus signifikante Verbesserungen von Lebenssinn, Hoffnung und spirituellem Wohlbefinden (Lo et al. 2014; Breitbart & Poppito 2014; Schulz-Kindermann 2018). Vier dieser Ansätze werden knapp skizziert, bevor der Hauptteil dieses Beitrags dem CALM-Ansatz gewidmet wird.

Existenzielle Psychotherapie (Yalom): Irvin Yaloms (1980) existenzielle Psychotherapie kann, neben der Logotherapie Viktor Frankls (2015), als eine wesentliche konzeptuelle Grundlage für sinnorientierte Therapieansätze gelten. Sie zielt auf vier Grundgegebenheiten mensch-

licher Existenz: Tod, Freiheit, Isolation und Sinnlosigkeit. Im Kontext schwerer Krebserkrankung gewinnen diese Dimensionen klinische Relevanz. Das Wissen um den eigenen Tod ist konkret gegenwärtig. Existenzielle Isolation im Angesicht des Todes – als letztlich nicht teilbare Erfahrung – kann im Zentrum des Erlebens Schwerstkranker und Sterbender stehen. Therapeutische Präsenz, das Aushalten von Ungewissheit und transparente, authentische Beziehungsgestaltung charakterisieren existenzielle Arbeit sensu Yalom.

Supportive-expressive Gruppentherapie (SEGT): David Spiegel und Yalom (1978) konzipierten gemeinsam einen gruppentherapeutischen Ansatz für schwerkranke Brustkrebspatientinnen, den Spiegel zur Supportive-Expressive Group Therapy (SEGT) weiterentwickelte (Spiegel & Classen 2000; Reuter & Spiegel 2016). Neben emotionalem Ausdruck und Erfahrungsaustausch hat das Programm den offenen Umgang mit Tod und Sterben zum Ziel. Randomisierte kontrollierte Studien (RCTs) zeigen, dass SEGT die psychische Belastung bei Patientinnen mit metastasierter Brustkrebserkrankung reduziert und die Lebensqualität signifikant verbessert (Vos et al. 2015).

Meaning-Centered Psychotherapy (Breitbart): Die Meaning-Centered Psychotherapy (MCGP) richtet sich an fortgeschritten erkrankte, hoch belastete Krebspatient:innen. Die acht Sitzungen umfassende MCGP verbindet psychoedukative Elemente mit erlebnisaktivierenden Ansätzen und Gruppenaustausch; theoretische Grundlage sind existenzphilosophische Überlegungen Frankls (2006; Breitbart & Poppito 2014). In RCTs erreichte MCGP gegenüber der Kontrollgruppe signifikant höhere Verbesserungen des spirituellen Wohlbefindens sowie eine Verringerung der psychischen Belastung (Breitbart et al. 2015). Der Ansatz ist auch als Einzeltherapie verfügbar (Individual Meaning-Centered Psychotherapy/IMCP; dt.: Mehnert et al. 2022).

Würdezentrierte Therapie/Dignity Therapy (Chochinov): Harvey Chochinov (2012) hat auf Basis umfangreicher Interviews mit Schwerstkranken das Erleben persönlicher Würde bzw. von Würdeverlust am Lebensende herausgearbeitet. Die Dignity Therapy (DT) richtet sich an Patient:innen in den letzten Lebenswochen und integriert Elemente der Biografiearbeit und des Life Review. Sie umfasst nur ein bis zwei Gespräche, deren Transkripte redaktionell bearbeitet und als persönliches Vermächtnis an Patient:innen und Angehörige übergeben werden (Chochinov et al. 2005). In einem RCT erzielte DT keine signifikante Reduktion psychischer Belastung, wurde aber von den Patient:innen selbst als hilfreich für das Würdeerleben bewertet (Chochinov et al. 2011).

Was die skizzierten Ansätze verbindet, ist die Betonung von Sinn, Selbsttranszendenz, Würde und Beziehung als Kernaspekten therapeutischer Arbeit. Was sie unterscheidet, sind Dauer, Zielgruppe und theoretische Verortung. Managing Cancer and Living Meaningfully (CALM; Nissim et al. 2011; Rodin et al. 2018) nimmt wesentliche Aspekte dieser Ansätze auf und entwickelt sie als einzeltherapeutische Kurzzeitintervention für Schwerstkranke weiter.

Managing Cancer and Living Meaningfully (CALM)

Konzept, Zielsetzung, theoretischer Hintergrund

Ein Team um Gary Rodin am Princess Margaret Hospital in Toronto entwickelte CALM auf dem Hintergrund umfangreicher empirischer Vorarbeiten und klinischer Praxis. CALM ist in mehreren psychologischen Schulen verwurzelt: in der relationalen Psychoanalyse (Mitchell 1988), in kognitiv-verhaltenstherapeutischen Ansätzen (Kissane et al. 1997), in der Bindungstheorie (Bowlby 1982) und in der existenziellen Therapie (Yalom 1980). CALM ist ein teilstrukturierter Kurzzeittherapie-Ansatz mit drei bis acht Sitzungen innerhalb eines halben Jahres, die zeitlich und inhaltlich flexibel an individuelle Verläufe und Befindlichkeiten schwer erkrankter Menschen angepasst werden.

Theoretische Grundlagen

Terror Management Theory. Die Terror Management Theory (Greenberg et al. 1986) geht davon aus, dass das Bewusstsein der eigenen Sterblichkeit eine existenzielle Grundangst erzeugt, die Menschen im Alltag durch kulturelle Sinnsysteme, Selbstwert und stabile Beziehungen abpuffern. Bei fortgeschrittener Krebserkrankung werden diese Schutzmechanismen erschüttert: Die Todesnähe lässt sich nicht mehr symbolisch distanzieren, und viele der bisherigen Quellen von Sinn, Identität und Zugehörigkeit geraten ins Wanken. CALM versteht diese Erschütterung weniger als Symptom, das behoben werden müsste, denn als Raum, in dem sich therapeutische Arbeit ereignet.

Bindungstheorie und sichere Basis. In existenziell bedrohlichen Situationen wird das Bindungssystem aktiviert und verstärkt das Bedürfnis nach Nähe und Schutz (Bowlby 1982). Individuelle Bindungsstrukturen können hilfreich und stützend, aber auch ängstlich und vermeidend sein. Die therapeutische Beziehung in CALM reflektiert diese Bindungstendenzen und kann in diesem Rahmen als sichere Basis erfahren werden – als Ort, von dem aus existenzielle Bedrohung innerlich repräsentiert werden kann, ohne zu überwältigen (Scheffold & Schulz-Kindermann 2022).

Mentalisierung. Das Konzept der Mentalisierung (Allen & Fonagy 2016) – die Fähigkeit, eigene und fremde mentale Zustände wahrzunehmen, zu verstehen und zu reflektieren, sowie psychologische Zustände von realen Fakten zu unterscheiden – spielt in CALM eine zentrale Rolle. Existenzielle Angst kann diese Fähigkeit erheblich einschränken und zu Vermeidung, Dissoziation oder affektiver Überflutung führen. CALM fördert die Mentalisierungsfähigkeit in der therapeutischen Beziehung und damit den inneren Abstand, der es erlaubt, die eigene Lage zu reflektieren, ohne von ihr überwältigt zu werden.

Doppelte Bewusstheit (Double Awareness). Ein spezifischer Beitrag von CALM ist das Konzept der Doppelten Bewusstheit (Rodin & Zimmermann 2008). Es beschreibt die Fähigkeit, zwei Bewusstseinsmodi zugleich zu halten: die Aufmerksamkeit für das gegenwärtige Leben – seine nach wie vor, vielleicht auch schmerzhaft, spürbare Ener-

gie, Beziehungen, Möglichkeiten – und das Bewusstsein der eigenen Endlichkeit und Sterblichkeit; keine starre Gleichzeitigkeit, sondern eine Pendelbewegung zwischen beiden Polen, je nach Anforderung der Situation. Das therapeutische Ziel ist nicht, die Endlichkeit zu verdrängen oder sie dauerhaft präsent zu halten, sondern die Kapazität zu entwickeln, sie – wenn nötig – zuzulassen und gleichwohl im Leben präsent zu bleiben (Colosimo et al. 2017).

In CALM wird eine vierdimensionale thematische Matrix entfaltet, die den therapeutischen Bearbeitungsraum strukturiert, ohne den Sitzungen eine starre Agenda aufzuzwingen (Rodin & Hales 2021): 1. *Medizinisch-pflegerische Fragen* wie die Kooperation mit medizinisch-pflegerischen Versorgern, Unterstützung bei anstehenden Entscheidungen, Optimierung von Behandlung und Symptomkontrolle; 2. *Das Selbst und die Anderen* mit Fragen des Identitätsgefühls und der krankheitsbezogenen Bindungsbeziehungen zu nahestehenden Menschen; 3. *Sinn, Ziele und Hoffnungen*, wie die Überprüfung der Quellen von Sinn im Leben und deren etwaige Neubewertung; 4. *Zukunft und Sterblichkeit* mit Fragen, die Ängste bzgl. des Sterbens, Sorgen über die Zukunft und spirituelle Themen angehen.

Diese vier Dimensionen sind kein Sitzungsplan, sondern eine inhaltliche «Landkarte», auf der thematische Gegenden eher in groben Zügen durchgemessen oder auch im Detail erkundet werden können. CALM versteht sich als systemischer Ansatz: Zugehörige Personen der Patient:innen sollen regelmässig zu Sitzungen eingeladen und in die therapeutischen Prozesse eingebunden werden. Strategisch vereint das therapeutische Vorgehen in CALM sowohl allgemein psychotherapeutische Elemente (z. B. bzgl. der therapeutischen Beziehung, Modulation von Affekten, Interpretation), psychoonkologische Ansätze (z. B. bzgl. Flexibilität, Einbezug, des sozialen Systems, Kontakt zu onkologischen Behandlern) und genuin CALM-spezifische Aspekte (wie z. B. Doppelte Bewusstheit oder Gestaltung der therapeutischen Beziehung als stabilisierendes Bindungsmodell).

CALM wurde in mehreren randomisierten kontrollierten Studien evaluiert, darunter eine multizentrische deutsche Studie (Mehnert et al. 2020; Scheffold et al. 2015). Die Befunde des kanadischen RCT (Lo et al., 2023) erbrachten, dass CALM-Teilnehmer:innen gegenüber der Kontrollbedingung signifikant stärkere Verbesserungen bei depressiven Symptomen, psychischem Wohlbefinden und der Vorbereitung auf das Sterben zeigten.

Therapeutische Haltung in der Arbeit mit Schwerstkranken

Therapeutisches Arbeiten mit fortgeschritten Erkrankten bedeutet Begleitung in Todesnähe. In dieser Grenzsituation befinden sich Psychotherapeut:innen in einem Dilemma: Sie sind trainiert, zur Stärkung des Selbst beizutragen. Todkranke ringen um den Erhalt dieses Selbst, während sie dabei sind, es zu verlieren. Todesangst kann mit Angst vor dem Nicht-mehr-Sein, vor dem letztend-

lichen Nichts verbunden sein. Weder der Tod selbst noch das sich dahinter auftuende Nichts sind (mit-)teilbare Erfahrungen. Ralf T. Vogel (2020) spricht hier von der «Opazität» dieser Phänomene, der Undurchschaubarkeit und letztendlichen Unaussprechlichkeit. Das ist für beide Seiten, die Erkrankten und die Therapeut:innen, schwer auszuhalten, insbesondere wenn zwei zentrale Bestrebungen beider berücksichtigt werden: die Sehnsucht der Totkranken nach Halt und das Bedürfnis der Therapeut:innen nach – möglichst verbal bearbeitbarer – Stabilisierung. Die therapeutische Haltung in CALM berücksichtigt dieses Dilemma. Sie ist geprägt von authentischer Präsenz, empathischem Verständnis und Offenheit für Ungewissheiten und Ambivalenzen. Giovanni Maio (2024) hat dazu in seiner Ethik der Verletzlichkeit auf eine grundlegende Wechselseitigkeit hingewiesen: Nicht nur die Kranken, sondern auch die Helfenden sind potenziell verletzlich. Maio sieht diese Reziprozität als wesentliche Voraussetzung allen therapeutischen Arbeitens.

Konkret heisst das: Therapeut:innen halten Impulse zur schnellen Einordnung zurück und lassen die Botschaften des Gegenübers zunächst durchdringen, bevor sie bewertet werden. Die therapeutische Beziehung wird nicht als «Wirkfaktor» instrumentalisiert, sondern als Möglichkeitsraum begriffen, in dem existenzielle Themen zwischen den Beteiligten entstehen können – weniger aktiv gestaltet als sich selbst entfaltend. Diese Haltung erfordert Ambiguitätstoleranz und die Fähigkeit, eigene Hilflosigkeit auszuhalten, ohne in therapeutischen Aktionismus zu verfallen. Gemeinsames Schweigen kann hier weniger als Hilflosigkeit, sondern als Anerkennung des letztlich Unbenennbaren und als ein wirkliches Miteinander Sein verstanden werden (Schulz-Kindermann 2020).

Klinisches Fallbeispiel aus der CALM-Arbeit¹

Die 23-jährige Patientin erkrankte im Alter von 16 Jahren an einem Ewing-Sarkom. Sie berichtet über diese Erkrankung recht nüchtern, lobt den damals hochgradig durchgeplanten Ablauf der Behandlung, mit eingreifender Operation ihres Unterschenkels und anschliessender systemischer Therapie. Ihre Zugehörigen – Eltern, Geschwister, Freund:innen – seien dagegen extrem belastet gewesen. Nach dem Abitur habe sie ein Studium begonnen, bis vor Kurzem ein ausgedehntes Rezidiv mit Metastasen in Lunge und Leber festgestellt worden sei. Sie wisse, wie bedrohlich diese Entwicklung sei, treffe aber weder in ihrem Elternhaus noch unter ihren Freund:innen jemanden, mit dem sie offen über diese Bedrohung sprechen könne. Es gehe alles viel zu schnell, ihre körperlichen Beschwerden nähmen praktisch täglich zu, unübersehbar schwellte ihr Aszites-Bauch nun mehr und mehr an. Sie suche eine vertrauensvolle Person, die vor dem Tabuthema Tod nicht zurückschreke.

Die Patientin sitzt von der Krankheit und der Behandlung sehr gezeichnet vor mir. Sie lächelt viel und hat junge, wache Augen, aber ihr Gang ins Praxiszimmer ist langsam und beschwerlich und ihr Bauch ist aufgrund von Aszites schmerzhaft geschwollen, die Atmung fällt ihr dadurch schwer. Bereits im ersten Gespräch in der Praxis lässt sich ein guter Kontakt mit der Ratsuchenden herstellen. Sie scheint tief erleichtert darüber, dass der Therapeut offenbar keine Scheu vor ihren erschreckenden und teilweise so widersprüchlichen Themen hat. Sie geniesst vielmehr das aufmerksame Interesse ihres Gegenübers: Wie sie denn mit dem gedunsenen Bauch umgehe, was sie mache, wenn die Todesängste sie überfallen, wie sie denn mit der Partnerin verfare, wenn diese ständig die kritischen Themen umschiffe etc. Sie merkt, wir können uns gemeinsam ihren kritischen Themen annähern, uns aber auch wieder entfernen. Meine eigene Haltung drückt andererseits auch Mitgefühl mit ihrer desolaten Situation und ein vorsichtiges Explorieren ihrer gegensätzlichen Emotionen aus, wobei dieses Gegensätzliche benannt und als anstrengend gekennzeichnet wird. Entscheidend erscheint meine authentische Präsenz und die Erlaubnis längerer Pausen, die die Bewegungen ihrer Gratwanderung mitzeichnet.

Nach diesen ersten Erkundungen schildere ich ihr kurz den therapeutischen Ansatz von CALM mit dem expliziten Hinweis, dass er speziell für Menschen mit einer fortgeschrittenen, palliativen Erkrankungssituation gedacht sei. Die Patientin irritiert das nicht, im Gegenteil, sie erlebt diese Einführung als Öffnung des Tabubereichs. Die vier Dimensionen des CALM-Ansatzes entfalten sich: Zunächst die Dimension der *medizinisch pflegerischen Themen*, denn die palliative onkologische Therapie läuft weiter bei einem sich rasant verschlechternden Allgemeinzustand und damit zunehmender Abhängigkeit von Anderen. Wir kommen schnell auf die Idee des Einschaltens eines SAPV-Teams (SAPV = Spezialisierte Ambulante Palliativversorgung), das sich der Bedürfnisse und insbesondere der zunehmenden Schmerzen annimmt. Dadurch verbessert sich der Zustand der Patientin sehr rasch – wenn auch nur vorübergehend. Daneben lernt sie in wenigen Tagen, Pflegende in Körperkontakt mit ihr zu lassen, denn die Erleichterung über die Abnahme der anstrengenden bis unmöglichen eigenen Körperpflege überwiegt. Gleichzeitig weist jede fremde Handlung an ihr auf die wachsende Schwäche und Abhängigkeit, also ihren nahenden Tod hin.

Herausfordernd ist die Dimension *Selbst und Andere* und zwar aus mehreren Gründen: Zunächst ist es die rasante Veränderung der Patientin selbst, die eine Vielzahl ihrer geschätzten Rollen ins Wanken und zum Einsturz bringt: Die Rolle als Partnerin, als Studierende, als selbstständig Lebende, als das Leben Geniessende ... Gleichzeitig erlebt sie die entsprechenden Reaktionen ihrer Gegenüber, die ihre Rollenveränderungen unterstreichen: Die Hilfesten, die sichtbare Angst und Sorge der anderen, aber auch die Zuwendung, das Zutrauen und die Kraft ihrer Nahestehenden. Ihre Rollen als Kind, Schwester und Partnerin sind unterschiedlich betroffen: Während die Patientin sich ihren Eltern gegenüber fallen lassen kann und zunehmend deren Versorgung akzeptiert,

¹ Dieses Fallbeispiel ist fiktional und aus der klinischen Praxis und der CALM-Therapie abgeleitet.

will sie mit ihren Freund:innen weiterhin auf Augenhöhe kommunizieren. Die therapeutische Strategie ist die wiederholter «Aufstellungen in sensu»: Da sie ohnehin die Begegnungen mit ihren Zugehörigen plastisch schildert und diese Vorstellungen «im Raum stehen», können diese befragt werden: Nach den jeweiligen vermuteten oder ausgesprochenen Motiven, nach ihren eigenen Bedürfnissen, v. a. nach dem Grundthema Alleinsein/Endgültig-alleinbleiben und In-Verbindung-Sein. Damit sind wir bei zwei Grundmotiven von CALM: Der Mentalisierung und der Beachtung des Bindungssystems. Ersteres scheint die Patientin mit einiger Leichtigkeit zu beherrschen: Sie hat ein recht sicheres Gespür für die Motivlagen ihrer Gegenüber, deren Ängste und Annäherungsbemühungen.

Gibt es in dieser Situation einen Platz für die Themen *Hoffnung und Sinn*, der dritten Dimension? Bei aller körperlichen und seelischen Gratwanderung erscheint sie nicht «demoralisiert», depressiv und hoffnungslos. Sie ist zerrissen zwischen den widerstrebenden Impulsen, hin- und hergerissen zwischen grossen und kleinen Zielen in der allernächsten Zeit und den unterschiedlichen Begegnungen mit ihren Nächsten. Sie hat täglich Hoffnung und sieht einen Sinn. Allerdings scheint die *Basis* für Sinn und Hoffnung ständig gefährdet. Das Pendeln zwischen diesen dynamischen Polen ist anstrengend, aber das (scheinbare) Fehlen eines übergeordneten Sinnes wohl eher hilfreich.

Die vierte Dimension, Themen um *Endlichkeit, Sterben, Tod und Vorsorge*, drängt häufig ins Gespräch. Es erscheint so, als hätte die Patientin in den vergangenen Wochen in einem Raum mit vielen Fenstern gesessen, die allesamt verschlossen und mit Fensterläden verdunkelt waren. Nun öffnet sich ein Fenster nach dem anderen und gibt Aussicht auf Bereiche, die fremd, unheimlich, aber auch traumhaft und schwerelos zu sein scheinen. Sie begibt sich in Gedanken und Anmerkungen bereits in eine jenseitige Position, wenn sie z. B. von sich in der Vergangenheit spricht oder davon, wie man sich an sie zurückerinnern wird. Der Kampf mit dem letzten Krankheitsfortschritt ist dabei einschneidend: Sie wird zunehmend müde, hat Mühe, sich zu konzentrieren und den Fokus zu halten. Immer häufiger muss sie sich den Nächsten ganz überlassen, was ihr zum Glück nicht so schwerfällt. Bei einem Hausbesuch zeigt sich die Patientin erfreut, ruhig und gefasst, wir wechseln wenige Worte, lassen gemeinsame Zeit verstreichen. Kurz danach stirbt sie, in relativer Ruhe, im Kreis ihrer Familie.

Ausblick

Psychotherapie bei fortgeschritten erkrankten Krebspatient:innen ist ein anspruchsvolles, aber sinnvolles klinisches Feld. Die hier dargestellten Ansätze zeigen, dass differenzierte und empirisch fundierte Konzepte für diese Patient:innengruppe entwickelt werden konnten. Im Mittelpunkt steht in allen Ansätzen die Arbeit mit Sinn, Würde, authentischer Beziehung und der Bereitschaft, sich der Realität von Endlichkeit zu stellen. CALM nimmt in diesem Feld eine besondere Stellung ein: Es ist manuali-

siert und zugleich hoch flexibel; auf robuste theoretische Fundamente gestützt und empirisch in RCTs belegt. Das Konzept der Doppelten Bewusstheit, die bindungstheoretische Rahmung und die Betonung eines geteilten echten Begegnungsraumes bilden den Kern einer therapeutischen Haltung, die für die Arbeit mit Schwerstkranken richtungsweisend ist. Die Integration existenzieller und palliativpsychotherapeutischer Konzepte in Aus- und Weiterbildung von Psychotherapeut:innen ist bislang unzureichend. Wer im ambulanten oder stationären Kontext Krebspatient:innen behandelt, wird früher oder später mit fortgeschrittener Erkrankung konfrontiert – und braucht dann nicht nur Methode, sondern Haltung und persönliche Ressourcen, um dieser Begegnung gerecht zu werden. Insbesondere im Rahmen der Weiterbildung Psychosoziale Onkologie (WPO), der Deutschen Gesellschaft für Palliativmedizin mit der Sektion Palliativpsychologie sowie des Instituts für Palliativpsychologie werden spezifische Fort- und Weiterbildungsmöglichkeiten angeboten.

Literatur

- Allen, J. G. & Fonagy, P. (2016). *Mentalisierungsgestützte Therapie: Das MBT-Handbuch. Konzepte und Praxis*. 3. Aufl. Klett-Cotta.
- Best, M. et al. (2015). Conceptual analysis of suffering in cancer: A systematic review. *Psycho-Oncology*, 24(9), 977–986.
- Bowlby, J. (1982). *Attachment and loss* (Vol. 1, 2nd ed.). Basic Books.
- Breitbart, W. & Poppito, S. (2014). *Individual meaning-centered psychotherapy for patients with advanced cancer: A treatment manual*. Oxford University Press. Deutsch: *Sinnzentrierte Einzel- und Gruppenpsychotherapie für Patienten mit fortgeschrittener Krebserkrankung: Ein Behandlungsmanual*. Übers. v. A. Mehnert et al.
- Breitbart, W. et al. (2000). Meaning-centered group psychotherapy for patients with advanced cancer: A pilot randomized controlled trial. *Psycho-Oncology*, 9(5), 358–367.
- Breitbart, W. et al. (2015). Meaning-centered group psychotherapy: An effective intervention for improving psychological well-being in patients with advanced cancer. *Journal of Clinical Oncology*, 33(7), 749–754.
- Boutros, A. et al. (2021). Health-related quality of life in cancer patients treated with immune checkpoint inhibitors in randomised controlled trials: A systematic review and meta-analysis. *European Journal of Cancer*, 159, 154–166.
- Cassell, E. J. (1999). Diagnosing suffering: A perspective. *Annals of Internal Medicine*, 131(7), 531–534.
- Chochinov, H. M. (2006). Dying, dignity, and new horizons in palliative end-of-life care. *Cancer Journal for Clinicians*, 56(2), 84–103.
- Chochinov, H. M. (2012). *Dignity therapy: Final words for final days*. Oxford UP.
- Chochinov, H. M. et al. (2005). Dignity therapy: A novel psychotherapeutic intervention for patients near the end of life. *Journal of Clinical Oncology*, 23(24), 5520–5525.
- Chochinov, H. M. et al. (2011). Effect of dignity therapy on distress and end-of-life experience in terminally ill patients: A randomised controlled trial. *Lancet Oncology*, 12(8), 753–762.
- Colosimo, K. et al. (2017). «Double awareness» in psychotherapy for patients living with advanced cancer. *Journal of Psychotherapy Integration*, 28(2), 125–140.
- Faller, H. et al. (2013). Effects of psycho-oncologic interventions on emotional distress and quality of life in adult patients with cancer: Systematic review and meta-analysis. *Journal of Clinical Oncology*, 31(6), 782–793.
- Frankl, V. E. (2006). *... trotzdem Ja zum Leben sagen*. 4. Aufl. Kösel.
- Frankl, V. E. (2015). *Grundkonzepte der Logotherapie*. Fakultas.

- Greenberg, J. et al. (1986). The causes and consequences of a need for self-esteem: A terror management theory. In R. F. Baumeister (Ed.), *Public self and private self* (S. 189–212). Springer.
- Kissane, D.W. et al. (1997). Cognitive-existential group therapy for patients with primary breast cancer – techniques and themes. *Psychooncology*, 6(1), 25–33.
- Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft, Deutsche Krebshilfe, AWMF) (2023). *Psychoonkologische Diagnostik, Beratung und Behandlung von erwachsenen Krebspatient*innen* (AWMF-Registernummer: 032-051OL). <https://www.leitlinienprogramm>
- Lo, C. et al. (2014). Managing Cancer And Living Meaningfully (CALM): Phase 2 trial of a brief individual psychotherapy for patients with advanced cancer. *Palliative Medicine*, 28(3), 234–242.
- Lo, C. et al. (2023). Efficacy of Managing Cancer And Living Meaningfully psychotherapy on depression, demoralization, and preparing for end of life in patients with advanced cancer. *Journal of Clinical Oncology*, 41(7), 1488–1499.
- Maio, G. (2024). *Ethik der Verletzlichkeit: Für eine Medizin der Zuwendung*. Herder.
- Marzorati, C. et al. (2025). Psychosocial factors associated with quality of life in cancer patients undergoing treatment: an umbrella review. *Health Qual Life Outcomes*, 23, 31.
- Mehnert, A. et al. (2020). Efficacy of the Managing Cancer and Living Meaningfully (CALM) individual psychotherapy for patients with advanced cancer: A single-blind randomized controlled trial. *Psycho-Oncology*, 29, 1895–1904.
- Mitchell, S.A. (1988). *Relational concepts in psychoanalysis: An integration*. Harvard UP.
- Nissim, R. et al. (2011). Managing Cancer And Living Meaningfully (CALM): A qualitative study of a brief individual psychotherapy for individuals with advanced cancer. *Palliative Medicine*, 26(5), 713–721.
- Philipp, R. et al. (2025). Existential distress in advanced cancer: A cohort study. *Gen Hosp Psychiatry*, 94, 184–191.
- Puchalski, C. et al. (2009). Improving the quality of spiritual care as a dimension of palliative care. *Journal of Palliative Medicine*, 12(10), 885–904.
- Reuter, K. & Spiegel, D. (2016). *Psychische Belastungen bei Krebserkrankungen. Gruppentherapie nach dem supportiv-expressiven Ansatz*. Hogrefe
- Rodin, G. & Hales, S. (2021). *Managing Cancer and Living Meaningfully. An Evidence Based Intervention for Cancer Patients and Their Caregivers*. Oxford UP.
- Rodin, G. & Zimmermann, C. (2008). Psychoanalytic reflections on mortality: A reconsideration. *Journal of the American Academy of Psychoanalysis and Dynamic Psychiatry*, 36(2), 181–196.
- Rodin, G. et al. (2018). Managing Cancer And Living Meaningfully (CALM): A randomized controlled trial of a psychological intervention for patients with advanced cancer. *Journal of Clinical Oncology*, 36(23), 2422–2432.
- Rodríguez-Prat, A. et al. (2024). The Wish to Hasten Death in Patients With Life-Limiting Conditions. A Systematic Overview. *Journal of Pain and Symptom Management*, 68, e91–e115
- Saunders, C. (1991). *Hospice and palliative care: An interdisciplinary approach*. Edward Arnold.
- Scheffold, K. & Schulz-Kindermann, F. (2022). Attachment in psychotherapeutic work with advanced cancer patients. A case study. *Verhaltenstherapie & Verhaltensmedizin*, 42(4), 299–310
- Scheffold, K. et al. (2015). Efficacy of a brief manualized intervention Managing Cancer And Living Meaningfully (CALM) adapted to German cancer care settings: Study protocol for a randomized controlled trial. *BMC Cancer*, 15, 592.
- Schulz-Kindermann, F. (2018). Existenzielles Leiden und sinnbezogene Interventionen. *Onkologische Welt*, 9(4)174–179
- Schulz-Kindermann, F. (2020). Authentische Präsenz in der Begegnung mit Schwerkranken. In *Psychoonkologie: Berührtsein zwischen Nähe und Distanz*. Pabst.
- Schulz-Kindermann, F. (2021). *Psychoonkologie. Grundlagen und Psychotherapeutische Praxis*. Beltz.
- Spiegel, D. & Classen, C. (2000). *Group therapy for cancer patients: A research-based handbook of psychosocial care*. Basic Books.
- Spiegel, D. & Yalom, I. (1978). «A support group for dying patients». *International Journal of Group Psychotherapy*, 28, 233–245
- Vehling, S. & Kissane, D.W. (2018). Existential distress in cancer. *Psycho-Oncology*, 27(11), 2525–2530.
- Vogel, R.T. (2020). *Existenzielle Themen in der Psychotherapie*. Kohlhammer
- Vos, P.J. et al. (2015). Psychosocial intervention for women with primary, non-metastatic breast cancer: A comparison between participants and non-participants. *Psycho-Oncology*, 15(3), 202–212.
- Wilson, K.G. et al. (2007). Suffering with advanced cancer. *Journal of Clinical Oncology*, 25(13), 1691–1697.
- Yalom, I.D. (1980). *Existential psychotherapy*. Basic Books.

Psychotherapy for cancer patients with advanced disease

Abstract: Advanced cancer represents not only a medical but an existential crisis. Beyond diagnosable psychiatric disorders – anxiety, depression, adjustment disorders – many patients experience a form of distress that is diagnostically more elusive: existential suffering, demoralization, and the loss of meaning and dignity. Modern oncological treatments have in many cases substantially extended life expectancy even in the palliative stage, giving rise to new clinical realities that have yet to be adequately addressed by psychotherapy. This article describes the theoretical and empirical framework of evidence-based psychotherapeutic approaches for patients with advanced cancer, with a focus on the CALM approach (Managing Cancer And Living Meaningfully): a manualized, meaning-oriented brief therapy derived from empirical and clinical evidence and grounded in multiple therapeutic traditions. Randomized controlled trials demonstrate its efficacy with respect to depressive symptoms, psychological well-being, and preparation for end of life. A detailed example from psycho-oncological practice outlines the psychotherapeutic work with CALM.

Keywords: psycho-oncology, advanced cancer, existential suffering, CALM, palliative psychotherapy, double awareness, dignity, meaning-making

Biografische Notiz

Dr. phil. Frank Schulz-Kindermann, Dipl.-Psych., ist Psychologischer Psychotherapeut (VT/GT), Psychoonkologe, Supervisor und arbeitete 1990–2021 am Universitätsklinikum Hamburg Eppendorf (UKE) in der Psychoonkologischen Versorgung, Lehre und Forschung. 2006–2021 war er Leiter des Psychoonkologischen Liaison- und Konsildienstes und der Spezialambulanz für Psychoonkologie am Institut für Medizinische Psychologie des UKE. Er war als klinischer Leiter an der multizentrischen Studie zu CALM (Managing Cancer and Living Meaningfully) beteiligt und unterrichtet dieses Verfahren weiterhin. Heute arbeitet er in privater psychotherapeutischer Praxis in Hamburg.

Kontakt

E-Mail: praxis-schulz-kindermann@posteo.de
Web: www.praxis-schulz-kindermann.de

Zur Modulation autonomer Regulationsprozesse bei ME/CFS innerhalb der Psychotherapie

Marie-Christine Hartlieb

Psychotherapie-Wissenschaft 16 (2) 2026 49–54

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-49>

Zusammenfassung: Der vorliegende Beitrag vertieft und erweitert bestehende psychotherapeutische Zugänge bei Myalgischer Enzephalomyelitis/Chronischem Fatigue-Syndrom (ME/CFS). Im Zentrum stehen Interventionen, die darauf abzielen, die Erholungsdynamik des autonomen Nervensystems (ANS) nach Belastungsreizen zu modulieren. Autonome Dysfunktionen sind ein häufiges Merkmal von ME/CFS. Studien zeigen im Vergleich zu Kontrollgruppen nicht nur erhöhte orthostatische Beschwerden und signifikante Veränderungen kardiovaskulärer Parameter, sondern auch autonome Dysfunktionen in unterschiedlichen Organsystemen. Darüber hinaus wurde eine reduzierte parasympathische Reaktivität nach Belastung nachgewiesen. Da klassische psychotherapeutische Ansätze, die ME/CFS primär als psychosoziale Stressreaktion interpretieren, von vielen Betroffenen als belastend erlebt wurden, bedarf es zwingend Anpassungen unter Berücksichtigung der pathophysiologischen Belastungsintoleranz. Bereits in einer früheren Publikation hat die Autorin erste Ansätze hierzu skizziert. Angesichts der aktuellen Forschungslage rückt sie die prolongierte autonome Anpassungsphase nach Belastung in den Vordergrund, die möglicherweise zur Persistenz von Fatigue und der Post-Exertional Malaise (PEM) beiträgt. PEM ist definiert als eine systemische Symptomverschlechterung nach Belastung, die oft zeitverzögert auftritt. Interventionen zur Reduktion der sympathischen Aktivierung werden vorgeschlagen. Es ist denkbar, dass eine gezielte psychotherapeutische Unterstützung bei der Regulation des ANS dazu beitragen könnte, die Erholungsprozesse nach Belastung und nach PEM zu verbessern. Das Therapieziel verlagert sich dabei von der Leistungssteigerung zur Verbesserung der parasympathischen Regenerationsfähigkeit. Bislang existieren keine empirischen Untersuchungen zu einem solchen psychotherapeutischen Zugang bei ME/CFS.

Schlüsselwörter: Myalgische Enzephalomyelitis/Chronisches Fatigue-Syndrom (ME/CFS), Psychotherapie, autonome Dysfunktion, Post-Exertional Malaise (PEM)

ME/CFS (Myalgische Enzephalomyelitis/Chronisches Fatigue-Syndrom) ist eine komplexe Multisystemerkrankung, die von der WHO bereits seit 1969 als neurologische Erkrankung unter dem ICD-Code G93.3 klassifiziert wird (Hoffmann et al., 2024; NICE, 2019). Im Zuge der Covid-19-Pandemie hat das Krankheitsbild eine erhöhte Beachtung erfahren, da ME/CFS als die schwerste Verlaufsform innerhalb des Post-COVID-Spektrums gilt. Das Leitsymptom ist die Post-Exertional Malaise (PEM), eine ausgeprägte, oft zeitverzögerte Erschöpfung und disproportionale Symptomverschlechterung sowie eine gestörte Erholungsdynamik nach physischer, kognitiver, sensorischer, orthostatischer oder emotionaler Belastung (Deutsche Gesellschaft für ME/CFS, 2026; Carruthers et al., 2011). Zur Diagnosesicherung von ME/CFS dienen international etablierte Kriterien wie die Kanadischen Konsenskriterien (CCC) (Hoffmann et al., 2024). Neben kognitiven Einschränkungen, Schmerzen und Schlafstörungen sind neuroimmunologische und orthostatische Beschwerden prägend. PEM ist dabei kein rein subjektives Phänomen, sondern lässt sich objektivieren, etwa durch eine 2-Tages-CPET-Untersuchung (Cardiopulmonary Exercise Testing) oder einen Kipptisch-Test (Moore et al., 2023; van Campen et al., 2021). Damit kann eine pathologische Beeinträchtigung der kardiopulmona-

len Erholungsfähigkeit nachgewiesen werden. Aktuelle Fragebögen wie bspw. der DSQ-PEM2 unterstützen die klinische Erfassung (Jason & Chee, 2025). Aufgrund der krankheitsbedingten Belastungsintoleranz stoßen konventionelle Behandlungs- und Psychotherapiekonzepte an ihre Grenzen (Grande et al., 2023; Hartlieb, 2021; Geraghty & Blease, 2019). Neben flexiblen Anpassungen wie individuell angepassten Intervallen, verkürzten Sitzungen oder der Möglichkeit von Online-Konsultationen zur Vermeidung von Anfahrtswegen, müssen auch die Therapieinhalte modifiziert werden. Eine ressourcenschonende Begleitung steht im Vordergrund.

Abgrenzung zu anderen Erkrankungen mit krankheitsspezifischer Fatigue

Fatigue kann bei psychischen Störungsbildern (wie Depressionen, Angststörungen, somatoformen Störungen) und organischen Erkrankungen (mit bspw. onkologischen, kardiologischen oder autoimmunologischen Ursachen) als Begleiterscheinung auftreten. Für diese Erkrankungen haben sich multimodale Therapiekonzepte auf Basis des biopsychosozialen Modells unter Einbezug somatischer und psychotherapeutischer Ansätze sowie aktivierender

Massnahmen bewährt. Zu den Inhalten zählt neben der somatischen Behandlung auch Psychoedukation zum Verständnis des Zusammenspiels von Körper, Psyche und Umwelt. Speziell für die psychotherapeutische Behandlung sind die Reduktion des Leidensdrucks, die Förderung von Emotions- und Stressregulation, der Abbau von Schon- und Vermeidungsverhalten sowie die Aktivitätssteigerung zentral. Darüber hinaus stehen die Stärkung von Selbstwirksamkeit, die kognitive Bearbeitung dysfunktionaler Denk- und Verhaltensmuster sowie die Aufarbeitung biografisch relevanter Belastungen und Beziehungserfahrungen im Fokus. Diese Aspekte werden in der Leitlinie *Funktionelle Körperbeschwerden* differenziert dargestellt und ausführlich erläutert (AWMF, 2018). Im Unterschied zu ME/CFS stellt PEM bei den genannten Erkrankungen jedoch kein zentrales diagnostisches Kernmerkmal dar.

In den deutschen Leitlinien *Müdigkeit* und *Long/Post-Covid* wird bei Patient*innen mit PEM ausdrücklich vor aktivierenden Massnahmen gewarnt (DEGAM, 2022; DGP et al., 2025). Die Kognitive Verhaltenstherapie (KVT) wird zwar als Angebot zur Krankheitsbewältigung empfohlen, ist jedoch nicht kurativ ausgerichtet. Früher wurde der KVT bei ME/CFS ein grosser Stellenwert zugeschrieben, da angenommen wurde, dass v. a. ungünstige Gedankenmuster und körperliche Dekonditionierung die Beschwerden aufrechterhalten würden. Diese Annahmen werden jedoch zunehmend kritisch bewertet (Weir & Speight, 2021; Grande et al., 2023). Das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen [IQWiG] (2023) nennt vorrangig Pacing (ein Energiemanagement-Tool, vgl. Goudsmit et al., 2011) als geeignete Strategie im Umgang mit der Erkrankung. Die KVT kann hilfreich sein, um zu lernen, mit den Grenzen der eingeschränkten Belastbarkeit, den Symptomen und mit den Folgen der Erkrankung wie Depressionen und Ängsten besser umzugehen, aber eine Fokussierung auf potenziell krankheitserhaltende Verhaltensweisen tritt in den Hintergrund. Die psychotherapeutische Herangehensweise bei ME/CFS unterscheidet sich somit von jener bei fatigueassoziierten Erkrankungen ohne PEM. Eine Nichtberücksichtigung von PEM würde genau jene krankheitserhaltenden und symptomverstärkenden Prozesse fördern, die man eigentlich reduzieren möchte. Pacing zielt darauf ab, wiederkehrende Push-Crash-Zyklen und PEM zu minimieren, und trägt dadurch langfristig zu einer Stabilisierung bei. Das Vorliegen einer PEM hat auch Konsequenzen für stationäre Behandlungen. Klassische aktivierende Rehabilitationskonzepte gelten bei PEM-Betroffenen als problematisch (Hammer et al., 2024; Schweizerische Gesellschaft für ME & CFS, 2024). Durch nicht ausreichend angepasste stationäre Massnahmen besteht das Risiko einer anhaltenden Zustandsverschlechterung infolge von Aktivierung und Überlastung. Ein Klinikalltag mit festen Essenszeiten, langen Wegstrecken und erhöhter Reiz- und Lärmbelastung kann bereits die Kraftreserven des gesamten Tages aufbrauchen. Das ambulante Setting stellt für die Mehrheit der Betroffenen wohl den sichersten Behandlungsrahmen dar, idealerweise in interdisziplinären Fachzentren – die

sich derzeit schrittweise etablieren. Auch komorbide oder reaktive psychische Störungsbilder sollten vorrangig in ambulanten Strukturen und durch eine ambulante Psychotherapie versorgt werden.

Aktueller Forschungsstand zu PEM und autonomer Dysfunktion

Exemplarisch für den aktuellen Forschungsstand zu PEM und autonomer Dysfunktion werden im Folgenden einige ausgewählte Studien kurz skizziert.

In der grossen MCAM-Analyse (Issa et al., 2025) ist eine autonome Dysfunktion fast durchgehend nachweisbar; 97 % der Betroffenen zeigen entsprechende Symptome, die stark mit funktionellen Einschränkungen korrelieren. Bei 90 % der ME/CFS-Patient*innen konnten van Campen et al. (2020) eine zerebrale Minderdurchblutung im Stehen nachweisen. In seiner Folgestudie wiesen die Forscher*innen nach, dass bei 37 % der Patient*innen das Herz nicht ausreichend mit einer Frequenzsteigerung auf die körperliche Belastung im Stehen reagiert (chronotrope Inkompetenz, van Campen et al., 2023). Die umfassende Analyse verschiedener physiologischer Marker durch Walitt et al. (2024) bestätigt die biologischen Fehlregulationen bei ME/CFS. Die Übersichtsarbeit von Ivanovska et al. (2025) hebt übereinstimmende pathologische Mechanismen zwischen ME/CFS und Long Covid hervor, wie eine autonome Dysfunktion, Immundysregulation, endotheliale Dysfunktion, gastrointestinale und mitochondriale Funktionsstörungen. 2-Tages-CPET-Studien zeigen mittels eines 2-tägigen Belastungstests eine deutlich verminderte Leistungsfähigkeit am zweiten Tag sowie eine unzureichende Erholung von Kreislauf und Nervensystem (Keller et al., 2024; Stevens et al., 2018; Moore et al., 2023). Habermann-Horstmeier und Horstmeier (2026) bilden vier Symptomcluster, von denen ein Cluster die autonome Dysfunktion darstellt. Die drei anderen Cluster werden den Bereichen Gehirn/ZNS, Immunsystem und Magen-Darm zugeordnet. Sogenannte *Brückensymptome* zeigen, wie die einzelnen Cluster miteinander verwoben sind. Die Auffassung, dass es sich um eine neuroimmunologische Multisystemerkrankung mit autonomer Dysfunktion handelt, wird erhärtet.

Neuere Arbeiten zu PEM belegen eindrücklich, dass PEM keine herkömmliche Erschöpfung darstellt, sondern eine biologische Grundlage hat: Germain et al. (2025) zeigen eine gravierende biologische Fehlanpassung nach Belastung, die sich durch anhaltende Störungen im Immunsystem, im mitochondrialen Energiestoffwechsel und in der neuromuskulären Regulation äussert. Appelman et al. (2024) gelingt ein direkter Nachweis von Gewebeschäden im Muskelgewebe nach körperlicher Anstrengung. Die Arbeitsgruppe um Lipkin (Che et al., 2025) identifiziert eine überschüssige Antwort des angeborenen Immunsystems als möglichen Auslöser von PEM. Wirth und Scheibenbogen (2020) beschreiben auf Basis ihrer Forschungen ein theoretisches Modell für die vaskuläre Dysregulation durch Störungen an adrenergen Rezeptoren, die die Gefässweite kontrollieren. Baraniuk

(2025) belegt einen gestörten Gehirnstoffwechsel bei ME/CFS durch Untersuchungen des Liquors.

Insgesamt verdeutlichen diese Befunde, dass PEM keine psychogene Reaktion darstellt. Die autonome Dysregulation stellt dabei ein zentrales Element dar (Issa et al., 2025). Während frühere Untersuchungen die orthostatische Intoleranz (OI) und das posturale Tachykardiesyndrom (POTS) oft als alleiniges Kriterium für eine autonome Dysfunktion werteten, zeigt die umfassende Erhebung der Symptome aller Organsysteme, dass die Dysautonomie nahezu die gesamte Patientenkohorte betrifft und nicht auf eine ME/CFS-Subgruppe beschränkt bleibt. Bei Issa et al. (2025) wird die autonome Dysfunktion klar als Kernmerkmal der Erkrankung herausgestellt.

Verlängerte autonome Erholungsphase

Bei gesunden Menschen steigt bei Aktivierung – etwa durch körperliche Anstrengung oder psychophysiologischen Stress – zunächst die innere Anspannung an (sympathische Aktivierung), normalisiert sich danach aber rasch über parasympathische Gegenregulation und zentrale regulatorische Mechanismen. Bei ME/CFS gibt es jedoch Hinweise auf eine verlängerte sympathische Aktivierung, während die beruhigenden Gegenmechanismen (parasympathische Gegenregulation) nur abgeschwächt einsetzen (Van Oosterwijck et al., 2021; Moore et al., 2023). Die Erholungsphase des ANS nach sympathischer Aktivierung dauert überproportional lange; die Rückregulation kardiovaskulärer und vegetativer Parameter auf ihr Ausgangsniveau ist entsprechend verlängert. Jede Form von Belastung – körperlich, geistig oder emotional – stellt eine physiologische Herausforderung für die Regulationsmechanismen ME/CFS-Betroffener dar. Die Autorin schlägt vor, der autonomen Dysfunktion im Rahmen der PEM verstärkte Aufmerksamkeit zu widmen und PEM nicht allein als Folge einer – häufig geringfügigen – Belastung zu betrachten, sondern auch als Ausdruck einer deutlich verlängerten autonomen Erholungszeit. Demnach benötigt das ANS nach körperlichen, kognitiven oder emotionalen Reizen unverhältnismässig lange, um zum Ausgangsniveau zurückzukehren. Wiederholte Reize ohne ausreichende parasympathische Rückregulation würden folglich zu einer anhaltend erhöhten autonomen Aktivierung führen. Eine chronische Überaktivierung des ANS bei gleichzeitig eingeschränkter parasympathischer Rückstellung kann in diesem Zusammenhang als Ausdruck einer erhöhten allostatischen Last interpretiert werden (Allostase-Konzept, vgl. Bobba-Alves et al., 2022). Das könnte bestehende Symptome verstärken.

Psychotherapie mit Fokus auf arousalregulierenden Prozessen

Psychotherapie gilt nicht als kurative Behandlung bei ME/CFS; sie kann zugrundeliegende pathophysiologische Veränderungen (bspw. Dysfunktionen der Mitochondrien)

nicht heilen, aber möglicherweise zur Unterstützung autonomer Regulationsprozesse beitragen. Da das ANS bei ME/CFS am Limit operiert, wirken intensive Emotionen oder kognitive Belastungen als zusätzliche Stressoren, welche die sympathische Überaktivierung verstärken. Psychotherapeutische Interventionen könnten die parasympathische Aktivität fördern, um die zusätzliche allostatische Last durch emotionale Trigger zu minimieren. Die Therapie wäre, im Rahmen dieses Ansatzes, somit auf Beruhigung und Verkürzung der Erholungsphase nach Reizen jeglicher Art ausgerichtet. Zentral könnte die Wahrnehmung von Übergängen zwischen Reiz und Nachstressreaktion sein: Patient*innen lernen, minimal erhöhte Aktivierungszustände früh zu erkennen und sofort regulierende Massnahmen einzuleiten. Dazu gehören unmittelbare Erdung nach Belastung, rechtzeitige Minipausen und Sich-Hinlegen, Atemregulation zur Förderung parasympathischer Aktivität und andere Techniken. Ziel ist es, Aktivierungsanstiege frühzeitig zu beenden. Das Ziel innerhalb jeder Therapieeinheit wäre, eine «sympathische Eskalation» bereits im Ansatz erkennen zu lernen, zu unterbrechen und anstatt sie auszuhalten oder zu vertiefen, möglichst frühzeitig zu beruhigen und zu regulieren. In diesem Modell läge der Fokus der Psychotherapie auf Stabilisierung, nicht auf dem vertiefenden Sich-Einlassen auf Erfahrungen und Gefühle. Ziele sind die Minimierung autonomer Überaktivierung und deren Nachwirkungen.

Möglichkeiten der therapeutischen Begleitung

Prävention & Pacing: Psychoedukation hinsichtlich des Krankheitsbildes und PEM, frühzeitige Vermittlung von Pacing-Strategien. Pacing gilt als der Goldstandard für ME/CFS (Goudsmit et al., 2011). Das Ziel ist, unter der individuellen Belastungsgrenze zu bleiben, um schwere und anhaltende Zustandsverschlechterungen zu vermeiden.

Emotionsregulation: Reduzierung emotionaler Trigger zur Entlastung des ANS. Selbst bei schwerstbetroffenen Patient*innen gibt es die Möglichkeit, emotionale Trigger zu entschärfen, die bspw. durch die häusliche Wohnsituation und die tägliche intensive Kommunikation mit Pflegepersonen entstehen können. Ein systemtherapeutisches Vorgehen, das im häuslichen bzw. familiären und ggf. schulischen Kontext ansetzt, könnte unterstützend wirken.

Co-Regulation: Beruhigende und stabilisierende Präsenz des*der Therapeuten*Therapeutin innerhalb der therapeutischen Beziehung.

Regulierende, körperbezogene Techniken – körperbezogene Erdungs- und Druckreize (z. B. Bodenkontakt, isometrische Muskelspannung, fester Halt): Deutliche Signale aus Muskeln und Gelenken werden vom Gehirn bevorzugt verarbeitet. Diese körperlichen Rückmeldungen wirken unmittelbar (Yu et al., 2024; Peláez & Taniguchi, 2016; Payne et al., 2015). Die Beruhigung geschieht nicht durch ressourcenintensive kognitive Funktionen, sondern durch einen direkten mechanischen Druck. Ein aktivierender Reiz (bspw. ein emotionsgeladener Gedanke) kann so durch eine sofortige Regulationsphase (bspw. durch

einen bewussten, festen Füsse-Boden-Kontakt) abgefangen werden, noch bevor eine starke Aktivierung entsteht. Ziel ist die Dämpfung der Reaktionsamplitude. Bei ME/CFS könnten vorsichtig sanfte taktile Reize ausprobiert werden. Körperübungen bergen jedoch stets das Risiko einer Überlastung und sind insbesondere für Schwerbetroffene nicht geeignet.

Interozeption (vgl. Farb et al., 2015) als *Werkzeug des Pacing*: Die Wahrnehmung von Innenzuständen dient dem rechtzeitigen Abbruch von Belastung. Das Ziel ist: «Ich spüre den Stressmodus frühzeitig und lege mich sofort hin.» Oder: «Wenn mein Herz rascher zu schlagen beginnt, stoppe ich und beginne mit ruhigen Atemzügen.»

Atemregulation und Entspannungstechniken: Eine langsame Atemfrequenz mit verlängerter Ausatmung kann die vagale Aktivität fördern und die Herzratenvariabilität (HRV) erhöhen (Gerritsen & Band, 2018). Das Ziel wäre: «Ich beruhige mein ANS, damit ich weniger Energie verbrauche.»

Musiktherapie könnte für mild betroffene Patient*innen einen wertvollen Ansatz darstellen, da das Medium Musik per se ein stressregulierendes Potenzial besitzt. Unter Berücksichtigung der individuellen Reiztoleranz können Rhythmus und Klang Entspannungsprozesse unterstützen – behutsam angeleitet von qualifizierten Musiktherapeut*innen.

Hinsichtlich der allgemeinen Besonderheiten in der psychotherapeutischen Arbeit mit ME/CFS-Patient*innen wird auf die Arbeit von Grande et al. (2023) als Basispublikation verwiesen.

Zusammenfassung und Ausblick

Die autonome Dysfunktion bei ME/CFS ist gut belegt. Einige Studien deuten darauf hin, dass eine verlängerte autonome Erholungszeit nach Reizen besteht. Nach Hypothese der Autorin können psychotherapeutische Interventionen den Prozess der autonomen Erholungsphase entweder unbeabsichtigt verstärken oder – bei geeigneter Gestaltung – regulierend verkürzen. Wiederholte Aktivierungen ohne ausreichende parasympathische Rückregulation könnten zu einer kumulativen Übererregung führen und PEM begünstigen und verstärken – ein Effekt, der möglichst vermieden werden sollte. Das Therapieziel würde sich von der Wiederherstellung der Leistungsfähigkeit hin zur Dämpfung der autonomen Überreaktivität und zur Förderung der parasympathischen Erholungskapazität verschieben. Die vorsichtige Regulation des ANS innerhalb der Psychotherapie könnte zu einer langfristigen, wenn auch begrenzten Stabilisierung des Systems beitragen. Es würde weniger Energie durch anhaltende Stressreaktionen gebunden – ein zentraler Aspekt bei ME/CFS, da die verfügbaren Energiere Ressourcen ohnehin stark eingeschränkt sind. Herkömmliche psychotherapeutische Methoden, die Aktivierung, Konfrontation und vertiefte emotionale Introspektion beinhalten, sind wenig hilfreich und kontraindiziert, solange keine ausreichende autonome Stabilisierung vorhanden ist.

Einschränkend ist anzumerken, dass die autonome Dysregulation lediglich einen Teilaspekt der PEM darstellt und darüber hinaus bislang keine Wirksamkeitsstudien zu dem von der Autorin vorgeschlagenen Ansatz vorliegen. Darüber hinaus handelt es sich bei Patient*innen mit ME/CFS sowohl hinsichtlich Ätiologie als auch Symptomatik um eine heterogene Gruppe. Möglicherweise profitieren insbesondere Patient*innen mit einer ausgeprägten Symptomatik aus dem Bereich der autonomen Dysfunktion von dem dargestellten Ansatz. Die Autorin regt an, die formulierte Hypothese einer möglichen Modifizierbarkeit der verlängerten autonomen Erholungszeit mittels psychotherapeutischer Interventionen im Rahmen zukünftiger Studien zu untersuchen.

Abschliessend ist noch festzuhalten, dass das vorgestellte Konzept rein symptomorientiert ist. Es dient dem Energiemanagement und der Stabilisierung des ANS, trifft jedoch keine Aussage hinsichtlich zugrunde liegender Ursachen und der Entstehung der Erkrankung.

Erklärung zur Nutzung von KI

Zur sprachlichen Glättung einzelner Textpassagen wurde KI (Gemini) eingesetzt. Die inhaltliche Verantwortung für sämtliche Formulierungen liegt vollständig bei der Autorin.

Literatur

- Appelman, B. et al. (2024). Muscle abnormalities worsen after post-exertional malaise in long COVID. *Nature Communications*, 15(1), 17. <https://doi.org/10.1038/s41467-023-44432-3>
- AWMF (2018). *S3 Leitlinie «Funktionelle Körperbeschwerden» (AWMF-Registernummer 051-001)*. https://register.awmf.org/assets/guidelines/051-001_S3_Funktionelle_Koerperbeschwerden_2018-11-abgelaufen.pdf
- Baraniuk, J.N. (2025). Cerebrospinal fluid metabolomics, lipidomics and serine pathway dysfunction in myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). *Scientific Reports*, 15(1), 7381. <https://doi.org/10.1038/s41598-025-91324-1>
- Bobba-Alves, N. et al. (2022). The energetic cost of allostasis and allostatic load. *Psychoneuroendocrinology*, 146(105951), 105951. <https://doi.org/10.1016/j.psyneuen.2022.105951>
- Carruthers, B.M. et al. (2011). Myalgic encephalomyelitis: International consensus criteria: Review: ME: Intl. Consensus Criteria. *Journal of Internal Medicine*, 270(4), 327–338. <https://doi.org/10.1111/j.1365-2796.2011.02428.x>
- Che, X. et al. (2025). Heightened innate immunity may trigger chronic inflammation, fatigue and post-exertional malaise in ME/CFS. *Npj Metabolic Health and Disease*, 3(1), 34. <https://doi.org/10.1038/s44324-025-00079-w>
- DEGAM (2022). *S3-Leitlinie Müdigkeit (AWMF-Register-Nr. 053-002)*. https://register.awmf.org/assets/guidelines/053-002_S3_Muedigkeit_2023-01_01.pdf
- Deutsche Gesellschaft für ME/CFS (2026). *Praxisleitfaden Myalgische Enzephalomyelitis/Chronisches Fatigue Syndrom (ME/CFS)*. <https://praxisleitfaden.mecfs.de/mecfs>
- DGP et al. (2025). *S1-Leitlinie Long-/Post-COVID (Registernummer 020-027)*. https://register.awmf.org/assets/guidelines/020-027_S1_Long-Post-Covid_2025-08-verlaengert_01.pdf

- Farb, N. et al. (2015). Interoception, contemplative practice, and health. *Frontiers in Psychology*, 6, 763. <https://doi.org/10.3389/fpsyg.2015.00763>
- Geraghty, K.J. & Blease, C. (2019). Myalgic encephalomyelitis/chronic fatigue syndrome and the biopsychosocial model: a review of patient harm and distress in the medical encounter. *Disability and Rehabilitation*, 41(25), 3092–3102. <https://doi.org/10.1080/09638288.2018.1481149>
- Germain, A. et al. (2025). Temporal dynamics of the plasma proteomic landscape reveals maladaptation in ME/CFS following exertion. *Molecular & Cellular Proteomics*, 24(12), 101467. <https://doi.org/10.1016/j.mcpro.2025.101467>
- Gerritsen, R.J.S. & Band, G.P.H. (2018). Breath of life: The respiratory vagal stimulation model of contemplative activity. *Frontiers in Human Neuroscience*, 12, 397. <https://doi.org/10.3389/fnhum.2018.00397>
- Goudsmit, E.M. et al. (2011). Pacing as a strategy to improve energy management in myalgic encephalomyelitis/chronic fatigue syndrome: a consensus document. *Disability and Rehabilitation*, 34(13), 1140–1147. <https://doi.org/10.3109/09638288.2011.635746>
- Grande, T. et al. (2023). The role of psychotherapy in the care of patients with myalgic encephalomyelitis/chronic fatigue syndrome. *Medicina*, 59(4), 719. <https://doi.org/10.3390/medicina59040719>
- Habermann-Horstmeier, L. & Horstmeier, L.M. (2026). Symptom clusters in ME/CFS reflect distinct neuroimmune and autonomic pathophysiological mechanisms: a translational model. *Journal of Translational Medicine*, 24(1). <https://doi.org/10.1186/s12967-026-08159-1>
- Hartlieb, M.-C. (2021). Psychotherapeutisch orientierte Behandlungen beim Chronic Fatigue Syndrome. *Psychotherapie-Wissenschaft*, 11(1), 51–56. <https://doi.org/10.30820/1664-9583-2021-1-51>
- Hoffmann, K. et al. (2024). Interdisziplinäres, kollaboratives D-A-CH Konsensus-Statement zur Diagnostik und Behandlung von Myalgischer Enzephalomyelitis/Chronischem Fatigue-Syndrom. *Wiener klinische Wochenschrift*, 136(Suppl. 5), 103–123. <https://doi.org/10.1007/s00508-024-02372-y>
- Hammer, S. et al. (2024). «Der Reha-Gedanke muss bei dieser Erkrankung völlig neu gedacht werden». Qualitative Ergebnisse einer Online-Befragung zu Erfahrungen Betroffener mit stationärer Rehabilitation bei Long/Post-COVID. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 188, 14–25. <https://doi.org/10.1016/j.zefq.2024.05.007>
- IQWiG (2023). *Myalgische Enzephalomyelitis/Chronisches Fatigue-Syndrom (ME/CFS)*. <https://www.gesundheitsinformation.de/mecfs-myalgische-encephalomyelitis-chronisches-fatigue-syndrom.html>
- Issa, A. et al. (2025). Autonomic dysfunction in myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS): Findings from the multi-site clinical assessment of ME/CFS (MCAM) study in the USA. *Journal of Clinical Medicine*, 14(17), 6269. <https://doi.org/10.3390/jcm14176269>
- Ivanovska, M. et al. (2025). Differential characteristics and comparison between Long-COVID syndrome and Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS). *Biomedicines*, 13(11), 2797. <https://doi.org/10.3390/biomedicines13112797>
- Jason, L.A. & Chee, K.J. (2025). A brief, comprehensive measure of post-exertional malaise. *Exploration of Neuroprotective Therapy*, 5, 1004116. <https://doi.org/10.37349/ent.2025.1004116>
- Keller, B. et al. (2024). Cardiopulmonary and metabolic responses during a 2-day CPET in myalgic encephalomyelitis/chronic fatigue syndrome: translating reduced oxygen consumption to impairment status to treatment considerations. *Journal of Translational Medicine*, 22(1), 627. <https://doi.org/10.1186/s12967-024-05410-5>
- Moore, G.E. et al. (2023). Recovery from exercise in persons with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). *Medicina*, 59(3), 571. <https://doi.org/10.3390/medicina59030571>
- NICE (2019). *Myalgic encephalomyelitis (or encephalopathy)/chronic fatigue syndrome: diagnosis and management*. <https://www.nice.org.uk/guidance/ng206/chapter/Recommendations>
- Payne, P. et al. (2015). Somatic experiencing: using interoception and proprioception as core elements of trauma therapy. *Frontiers in Psychology*, 6, 93. <https://doi.org/10.3389/fpsyg.2015.00093>
- Peláez, F.J. & Taniguchi, S. (2016). The gate theory of pain revisited: Modeling different pain conditions with a parsimonious neurocomputational model. *Neural Plasticity*, 4131395. <https://doi.org/10.1155/2016/4131395>
- Schweizerische Gesellschaft für ME & CFS (2024). *Mitteilung zur Reha-Umfrage 2024*. https://long-covid-info.ch/wp-content/uploads/2024/12/Bieri_Mitteilung-zur-Reha-Umfrage-2024-2.pdf
- Stevens, S. et al. (2018). Cardiopulmonary exercise test methodology for assessing exertion intolerance in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Frontiers in Pediatrics*, 6, 242. <https://doi.org/10.3389/fped.2018.00242>
- van Campen, C.L.M.C. et al. (2020). Cerebral blood flow is reduced in ME/CFS during head-up tilt testing even in the absence of hypotension or tachycardia: A quantitative, controlled study using Doppler echography. *Clinical Neurophysiology Practice*, 5, 50–58. <https://doi.org/10.1016/j.cnp.2020.01.003>
- van Campen, C.L.M.C. et al. (2023). Orthostatic chronotropic incompetence in patients with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). *IBRO Neuroscience Reports*, 15, 1–10. <https://doi.org/10.1016/j.ibneur.2023.04.005>
- Van Oosterwijck, J. et al. (2021). Reduced parasympathetic reactivation during recovery from exercise in Myalgic Encephalomyelitis/Chronic Fatigue Syndrome. *Journal of Clinical Medicine*, 10(19), 4527. <https://doi.org/10.3390/jcm10194527>
- Walitt, B. et al. (2024). Deep phenotyping of post-infectious myalgic encephalomyelitis/chronic fatigue syndrome. *Nature Communications*, 15(1), 907. <https://doi.org/10.1038/s41467-024-45107-3>
- Weir, W. & Speight, N. (2021). ME/CFS: Past, present and future. *Healthcare*, 9(8), 984. <https://doi.org/10.3390/healthcare9080984>
- Wirth, K. & Scheibenbogen, C. (2020). A Unifying Hypothesis of the Pathophysiology of Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS): Recognitions from the finding of autoantibodies against β_2 -adrenergic receptors. *Autoimmunity Reviews*, 19(6), 102527. <https://doi.org/10.1016/j.autrev.2020.102527>
- Yu, J. et al. (2024). The effect of weighted blankets on sleep and related disorders: a brief review. *Frontiers in Psychiatry*, 15, 1333015. <https://doi.org/10.3389/fpsyg.2024.1333015>

Modulation of autonomic regulatory processes in ME/CFS within psychotherapy

Abstract: This article deepens and expands existing psychotherapeutic approaches to myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). The focus is on interventions that aim to modulate the recovery dynamics of the autonomic nervous system (ANS) after stress stimuli. Autonomic dysfunctions are a common feature of ME/CFS. Studies show not only increased orthostatic symptoms and significant changes in cardiovascular parameters compared to control groups, but also autonomic dysfunctions in various organ systems. In addition, reduced parasympathetic reactivity after stress has been demonstrated. Since classic psychotherapeutic approaches, which primarily interpret ME/CFS as a psychosocial stress reaction, have been experienced as stressful by many affected individuals, adjustments are required that take pathophysiological exercise intolerance into account. The author has already outlined some initial approaches to this in an earlier publication. In view of the current state of research, the author focuses on the prolonged autonomic readjustment phase after exertion, which may contribute to the persistence of fatigue and post-exertional malaise (PEM). PEM is defined as a systemic worsening of symptoms after exertion,

which often occurs with a temporal delay. Psychotherapeutic interventions for reducing sympathetic activity are proposed. It is possible that specific psychotherapeutic support in regulating the ANS could help to improve recovery processes after exertion and after PEM. The therapeutic goal shifts from improving performance to improving parasympathetic regeneration capacity. So far, there are no empirical studies on such a psychotherapeutic approach to ME/CFS.

Keywords: myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), psychotherapy, autonomic dysfunction, post-exertional malaise (PEM)

Biografische Notiz

Marie-Christine Hartlieb ist Absolventin des Lehrgangs für Musiktherapie an der Musikhochschule Wien (heutige Universität für Musik und darstellende Kunst) und war als Musiktherapeutin und Psychotherapeutin in Ausbildung unter Supervision in unterschiedlichen Arbeitsfeldern (Kinder und Jugendliche, Geriatrie, Psychiatrie, intellektuelle Behinderung) tätig.

Kontakt

E-Mail: c.hartlieb@gmx.at

Zur epistemischen Eigenständigkeit der Psychotherapiewissenschaft

Plädoyer für eine Ethik der reflexiven Erfahrung

Tamara M. Trebes

Psychotherapie-Wissenschaft 16 (2) 2026 55–61

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-55>

Zusammenfassung: Die zunehmende Akademisierung der Psychotherapie wirft grundlegende Fragen nach ihrer wissenschaftlichen Verortung und ethischen Fundierung auf. Insbesondere Tendenzen zur Standardisierung, Vereinheitlichung und Psychologisierung drohen jene Dimensionen zu marginalisieren, die für psychotherapeutische Praxis konstitutiv sind: subjektives Erleben, Prozessualität, Ambiguität und das Unauflösbare menschlichen Leidens. Der vorliegende Beitrag argumentiert, dass Psychotherapie einer eigenständigen epistemischen Logik folgt, die sich nicht in erklärenden oder technizistisch verkürzten Wissenschaftsmodellen aufgehen lässt. Vor diesem Hintergrund entwickelt der Artikel ein Plädoyer für eine Ethik der reflexiven Erfahrung als zentrale Voraussetzung der epistemischen Eigenständigkeit der Psychotherapiewissenschaft. Wissen wird dabei nicht als primäre Instanz verstanden, sondern als abgeleitete Kategorie, die sich erst im Durchgang durch reflexiv erschlossene Erfahrung, Beziehung und Zeit legitimiert. Lehrtherapie und Supervision werden als zentrale Erfahrungsräume ausgewiesen, in denen diese ethische Haltung ausgebildet und gegen epistemische Hybris abgesichert wird. Der Text versteht sich als Beitrag zur psychotherapiewissenschaftlichen Selbstverständigung und als Gegenposition zu vereinheitlichenden Tendenzen im aktuellen wissenschaftlichen Diskurs.

Schlüsselwörter: Psychotherapiewissenschaft, Ethik der reflexiven Erfahrung, epistemische Eigenständigkeit, subjektives Erleben, Intersubjektivität, Nicht-Wissen, Ambiguitätstoleranz, Unauflösbarkeit menschlichen Leidens, Psychotherapie

Psychotherapie befindet sich in einem tiefgreifenden Umbruch. Mit der zunehmenden Akademisierung ihrer Ausbildung rücken Fragen der Standardisierung, der Kompetenzorientierung und der Integration therapeutischer Verfahren verstärkt in den Vordergrund. Dabei bleibt häufig unhinterfragt, welches Verständnis von Kompetenz diesen Entwicklungen zugrunde liegt und welche Formen therapeutischen Könnens dadurch sichtbar – oder unsichtbar – werden. Was als Fortschritt erscheint – mehr Wissenschaftlichkeit, Vergleichbarkeit und Struktur –, geht mit einer Verschiebung einher, die selten explizit reflektiert wird: der stillen Veränderung dessen, was als legitim, wirksam und anschlussfähig gilt. Psychotherapie läuft Gefahr, zunehmend an Wissenschaftslogiken angepasst zu werden, die ihrem Gegenstand nur teilweise gerecht werden. Besonders deutlich zeigt sich diese Spannung dort, wo Integration nicht mehr als offene Suchbewegung verstanden wird, sondern als Annäherung an ein vereinheitlichendes Modell. Was ursprünglich die Exploration unterschiedlicher Menschenbilder, Methoden und Beziehungskonzepte ermöglichte, verwandelt sich zunehmend in den Anspruch auf Kohärenz, Effizienz und Manualisierbarkeit. Damit gerät jedoch genau das unter Druck, was Psychotherapie im Kern ausmacht: die Anerkennung subjektiver Wirklichkeit, die Offenheit für Ambivalenz und die Bereitschaft, dem Unauflösbaren Raum zu geben.

Psychotherapeutische Praxis und ihre wissenschaftliche Reflexion unterscheiden sich in ihrem Erkenntnisinteresse grundlegend von angrenzenden Disziplinen wie der Psychologie oder der Medizin. Während Psychotherapie auf Beziehung, Verstehen und Begleitung im therapeutischen Prozess zielt, richtet sich Psychotherapiewissenschaft auf die reflexive Untersuchung dieser Praxis und ihrer erkenntnistheoretischen Voraussetzungen. Ihr Gegenstand ist nicht das isolierte Verhalten oder die messbare Variable, sondern der Mensch in seiner biografischen, affektiven und intersubjektiven Verfasstheit. Diese Komplexität lässt sich nicht verlustfrei standardisieren, ohne das Subjektive zu verfehlen, das Psychotherapie überhaupt erst notwendig macht. Vor diesem Hintergrund stellt sich die Frage nach der ethischen Grundlage psychotherapeutischer Praxis neu. Der vorliegende Beitrag schlägt den Begriff einer *Ethik der reflexiven Erfahrung* vor, um jene Haltung zu beschreiben, die nicht aus normativen Vorgaben oder technischem Wissen erwächst, sondern aus dem gelebten Durchgang durch eigene psychische Prozesse. Lehrtherapie und Supervision erscheinen in diesem Verständnis nicht als ergänzende Ausbildungselemente, sondern als zentrale Träger einer epistemischen Ethik, die Demut, Beziehungsoffenheit und die Achtung vor dem Unauflösbaren schützt. In einer Zeit zunehmender Vereinheitlichung versteht sich dieser Ansatz als Plädoyer für eine Psychotherapiewissenschaft, die Vielfalt

nicht als Mangel, sondern als Konsequenz menschlicher Komplexität begreift.

Zur epistemischen Eigenständigkeit

Die Frage nach der wissenschaftlichen Verortung der Psychotherapie ist keine bloss institutionelle oder universitätspolitische, sondern eine erkenntnistheoretische. Mit der zunehmenden Einbettung psychotherapeutischer Ausbildung in universitäre Strukturen wird häufig implizit vorausgesetzt, dass Psychotherapie innerhalb psychologischer oder medizinischer Wissenschaftslogiken angemessen aufgehoben sei. Diese Annahme verkennt jedoch, dass Psychotherapie einem eigenständigen Erkenntnisinteresse folgt und einen anderen Gegenstandsbereich adressiert (Greiner, 2015, 2025).

Bereits Gerhard Benetka weist darauf hin, dass die moderne Psychologie historisch in einem Kontext entstand, in dem subjektive und sinnbezogene Dimensionen menschlichen Erlebens zunehmend aus dem wissenschaftlichen Denken ausgeschlossen wurden (Slunecko & Benetka, 2023, S. 10). Daraus ergibt sich die erkenntnistheoretische Frage, inwiefern psychotherapeutisches Verstehen – das sich zentral auf subjektives Erleben, Sinnzusammenhänge und relationale Erfahrung bezieht – bruchlos innerhalb psychologischer Erkenntnislogiken verortet werden kann, ohne dabei wesentliche Aspekte seines Gegenstands zu verlieren. Die Anerkennung epistemischer Differenzen bedeutet dabei nicht die Ablehnung interdisziplinärer Zusammenarbeit, sondern die Notwendigkeit, unterschiedliche wissenschaftliche Erkenntnisinteressen reflexiv auseinanderzuhalten. In jüngeren systematischen Arbeiten zur Psychotherapiewissenschaft wird diese Differenz ausdrücklich hervorgehoben. So beschreibt Paolo Raile (2023) Psychotherapiewissenschaft als eigenständige Disziplin, deren Erkenntnisinteresse nicht in Erklärung oder Prognose aufgeht, sondern auf das verstehende Arbeiten mit subjektivem Erleben, Beziehung und Prozessualität gerichtet ist.

Während Psychologie primär auf Erklärung, Klassifikation und Generalisierung zielt, richtet sich Psychotherapie auf subjektives Erleben in seiner biografischen, affektiven und relationalen Verfasstheit. Ihr Gegenstand ist nicht das isolierbare Merkmal oder das beobachtbare Verhalten, sondern der im therapeutischen Prozess hergestellte sinnhaft strukturierte Zusammenhang von Erleben, Ausdruck und Beziehung. Psychotherapeutisches Verstehen vollzieht sich damit nicht im Modus objektivierender Analyse, sondern als dialogischer, zeitlich entfalteter Prozess, der wesentlich an Sprache, Symbolisierung und Resonanz gebunden ist. In diesem Sinn wird psychotherapeutische Praxis in der Psychotherapiewissenschaft als eine Form des verstehenden Umgangs mit verbalen und nonverbalen Artikulationen subjektiven Erlebens gefasst (Greiner, 2024a, S. 9f.).

In mehreren Arbeiten hat Kurt Greiner (2015; Jandl & Greiner, 2023) darauf hingewiesen, dass Psychotherapie nicht als Unterform psychologischer Erkenntnis missver-

standen werden darf, sondern einer eigenen wissenschaftlichen Logik folgt, die hermeneutisch fundiert und auf Sinnbildung ausgerichtet ist. In der psychotherapeutischen Praxis konkretisiert sich diese Sinnbildung im relationalen Vollzug, d.h. im intersubjektiven Geschehen zwischen Therapeut*in und Patient*in. Psychotherapie erscheint in dieser Perspektive nicht als Anwendung vorgefertigter Theorien oder Techniken, sondern als eigenständige, schulenübergreifende Praxis, die den therapeutischen Prozess als offen und nicht vollständig vorhersagbar begreift.

Vor diesem Hintergrund wird deutlich, dass eine Psychologisierung der Psychotherapie nicht nur eine institutionelle Verschiebung darstellt, sondern mit einer inhaltlichen Reduktion einhergeht. Wird psychotherapeutisches Handeln primär über Wirksamkeitsnachweise, manualisierte Interventionen oder standardisierte Kompetenzmodelle definiert, geraten zentrale Dimensionen des therapeutischen Geschehens aus dem Blick. Greiner (2022) hat in diesem Zusammenhang kritisch herausgearbeitet, dass Wirksamkeitsstudien für sich genommen nichts über den Wissenschaftlichkeitsstatus psychotherapeutischer Modalitäten aussagen können, da sie zentrale erkenntnistheoretische Voraussetzungen psychotherapeutischer Praxis verfehlen. Die Forderung nach einer eigenständigen Psychotherapiewissenschaft ist daher nicht als blosse Abgrenzungsgeste zu verstehen, sondern als Versuch, unterschiedliche wissenschaftliche Erkenntnisinteressen differenziert zu reflektieren. Sie trägt dem Umstand Rechnung, dass menschliches Leiden nicht vollständig erklärbar, auflösbar oder normierbar ist, sondern einer Praxis bedarf, die Beziehung, Zeit und Ungewissheit ernst nimmt. In dieser Differenz zwischen Erklären und Verstehen liegt eine zentrale Voraussetzung dafür, Psychotherapie als eigenständige wissenschaftliche und zugleich ethisch verantwortliche Praxis zu begreifen.

Diversität als Konsequenz, nicht als Defizit

Die Vielfalt psychotherapeutischer Schulen und Methoden wird im wissenschaftlichen Diskurs häufig als Problem markiert: als Zeichen mangelnder Reife, unzureichender Konsolidierung oder fehlender Integration. Implizit steht dahinter die aus wissenschaftstheoretischen Paradigmenmodellen bekannte Annahme, dass wissenschaftlicher Fortschritt langfristig auf Vereinheitlichung hinauslaufe (vgl. Kuhn, 1976) – eine Vorstellung, die sich sowohl in integrativen und schulenübergreifenden Psychotherapiemodellen (vgl. Grawe et al., 1994; Fiedler, 2012) als auch in aktuellen Bestrebungen nach einer «unifying, science-based view of psychotherapy» (Smith, 2025) wiederfindet. Die Frage nach einer eigenständigen Psychotherapiewissenschaft ist innerhalb des deutschsprachigen Diskurses dabei keineswegs neu. Bereits frühere Beiträge verwiesen auf die Gefahr einer Reduktion psychotherapeutischer Praxis auf psychologische oder rein naturwissenschaftliche Modelle und betonten die Notwendigkeit einer eigenständigen epistemischen Verortung (vgl. Pritz & Teufelhart, 1996; Fischer et al., 2009). Vor dem Hin-

tergrund aktueller Akademisierungsprozesse und der zunehmenden institutionellen Anbindung psychotherapeutischer Ausbildung an psychologische Fakultäten gewinnt diese Diskussion jedoch erneut an Aktualität.

Nicht nur für die Psychotherapie greift dieses Fortschrittsverständnis jedoch zu kurz. Ihre Pluralität ist kein vorläufiger Zustand, sondern eine Konsequenz ihres Gegenstands. Psychotherapie arbeitet mit Menschen in ihrer je spezifischen Lebensgeschichte, Körperlichkeit, Affektivität und Beziehungserfahrung. Diese Dimensionen sind weder vollständig erfassbar noch widerspruchsfrei integrierbar. Unterschiedliche psychotherapeutische Schulen repräsentieren daher weniger konkurrierende Erklärungsmodelle als vielmehr unterschiedliche epistemische Zugänge zu menschlicher Erfahrung. Ihre Koexistenz verweist auf die Unabschliessbarkeit des Psychischen selbst. Versuche, psychotherapeutische Vielfalt im Sinne theoretischer oder methodischer Vereinheitlichung aufzulösen, laufen Gefahr, mit Reduktionen komplexer subjektiver Erfahrungsdimensionen einherzugehen. Integration wird dabei nicht als dialogische Suchbewegung verstanden, sondern als Angleichung an ein dominantes Paradigma, das festlegt, welche Formen des Verstehens, Behandelns und Wirkens als legitim gelten. Dort, wo Integration primär über Standardisierung und Vereinheitlichung organisiert wird, gerät psychotherapeutische Praxis zunehmend in Gefahr, primär über Techniken, Manuale und standardisierte Interventionen definiert, während subjektive, situative und relationale Dimensionen an den Rand gedrängt werden.

Diese Tendenz zur Vereinheitlichung lässt sich jedoch nicht nur innerhalb psychotherapeutischer Diskurse beobachten, sondern steht in einem weiteren kulturtheoretischen Zusammenhang. In seinem Essay *Die Vereindeutigung der Welt* beschreibt Thomas Bauer (2018) einen umfassenden Verlust an Ambiguitätstoleranz, der moderne Gesellschaften dazu verleitet, Mehrdeutigkeit zugunsten von Eindeutigkeit, Kontrolle und funktionaler Klarheit zurückzudrängen. Was nicht eindeutig, messbar oder widerspruchsfrei ist, gerät unter Rechtfertigungsdruck oder wird als defizitär markiert. Vor diesem Hintergrund erscheint der Wunsch nach einer *Einheitspsychotherapie*, wie er etwa im Kontext integrativer Modelle formuliert wurde (vgl. Grawe et al.), nicht nur als fachinterne Vereinfachung, sondern als Ausdruck einer breiteren kulturellen Bewegung, die Ambivalenz schwer erträgt. Psychotherapie widersetzt sich dieser Bewegung dort, wo sie Mehrdeutigkeit nicht als Mangel, sondern als konstitutives Merkmal menschlicher Erfahrung begreift. Ihre Pluralität wird damit zu einer Form epistemischer und ethischer Gegenwehr gegen die *Vereindeutigung* (vgl. Bauer, 2018) des Psychischen.

Eine solche Vereinheitlichung erkennt, dass psychotherapeutische Arbeit gerade dort ansetzt, wo Eindeutigkeit nicht herstellbar ist. Leiden folgt selten klaren Kausalitäten, Veränderung verläuft nicht linear, und Verstehen ist kein einmaliger Akt, sondern ein zeitlich offener, relationaler Prozess. Die Stärke psychotherapeutischer Praxis liegt nicht in der Reduktion dieser Komplexität,

sondern in der Fähigkeit, sie auszuhalten und fruchtbar zu machen. Aus psychotherapiewissenschaftlicher Perspektive ist Diversität daher nicht als mangelnde Systematik zu bewerten, sondern als epistemische Notwendigkeit. In der psychotherapiewissenschaftlichen Arbeit von Greiner lässt sich diese Vieldeutigkeit nicht als Defizit, sondern als Qualitätsmerkmal verstehen. Greiner (2017) zeigt, dass jene Eigenschaften, die aus naturwissenschaftlicher Perspektive als methodische Schwächen gelten – Kontextsensibilität, interpretative Offenheit und die Abwesenheit eindeutiger Zuordnungen –, im hermeneutisch-textwissenschaftlichen Verständnis gerade die Voraussetzung wissenschaftlicher Angemessenheit bilden. Unterschiedliche Schulen halten unterschiedliche Aspekte menschlicher Erfahrung im Blick, entwickeln unterschiedliche Sprachen für das Nicht-Sagbare und unterschiedliche Formen therapeutischer Beziehung. Diese Vielfalt wirkt als Schutz vor der Verabsolutierung einzelner Menschenbilder und vor dem Anspruch, zu wissen, was für einen Menschen richtig ist.

Integration bedeutet in diesem Sinn nicht Vereinheitlichung, sondern Dialogfähigkeit. Sie setzt die Bereitschaft voraus, Differenz auszuhalten – zwischen Theorien, Schulen und Praxen ebenso wie im intersubjektiven Raum der Therapie selbst. Eine Psychotherapie, die ihre Vielfalt zugunsten standardisierter Lösungen aufgibt, verliert nicht nur an theoretischer Tiefe, sondern an ethischer Sensibilität. Denn dort, wo das Subjektive dem Machbaren geopfert wird, beginnt die Gefahr des Übergriffs. In diesem Sinne lässt sich psychotherapiewissenschaftliche Reflexion als strukturell verwandt mit psychotherapeutischer Praxis verstehen. So wie Psychotherapie darauf zielt, implizite, unbewusste oder vorreflexive Anteile des Erlebens zugänglich und reflektierbar zu machen, richtet sich Psychotherapiewissenschaft darauf, die impliziten Voraussetzungen, Vorannahmen und Menschenbilder psychotherapeutischer Theorien sichtbar zu machen. Greiners (2015, 2024a) experimentell-psychotherapiewissenschaftlicher Zugang versteht Theorien nicht als abgeschlossene Systeme, sondern als historisch, kulturell und erkenntnistheoretisch situierte Konstruktionen, deren Wissenschaftlichkeit sich nicht allein an ihrer Funktionalität, sondern wesentlich an ihrer kritischen Reflexionsfähigkeit bemisst.

Psychotherapiewissenschaft wiederholt damit auf einer anderen Ebene das, was psychotherapeutische Arbeit selbst auszeichnet: Sie macht Implizites explizit, hält Spannungen aus und widersetzt sich der *Vereindeutigung* komplexer Prozesse. Wissenschaftlichkeit zeigt sich hier nicht in der Reduktion von Mehrdeutigkeit, sondern in der Fähigkeit, diese reflexiv zu bearbeiten. Diversität erscheint damit nicht als methodisches Defizit, sondern als notwendige Folge eines Erkenntnisinteresses, das dem Psychischen in seiner Prozessualität und Unabschliessbarkeit gerecht werden will. Diese strukturelle Nähe von Praxis und Wissenschaft verweist darauf, dass psychotherapiewissenschaftliche Reflexion eine ethische Haltung voraussetzt, die auf Verstehen zielt, dabei jedoch die Begrenztheit und Vorläufigkeit dieses Verstehens mitreflektiert, anstatt sie

zu verleugnen. Diese Haltung stellt keine Relativierung von Erkenntnis dar, sondern bildet die Bedingung verantwortlichen psychotherapeutischen Handelns, indem sie Verstehen an Beziehung, Kontext und Prozess bindet.

Wissen als abgeleitete Kategorie

Die bisherigen Überlegungen führen an einen Punkt, an dem die ethische Dimension psychotherapeutischer Praxis nicht länger über Normen, Richtlinien oder professionelle Selbstbeschreibungen bestimmt werden kann. Stattdessen rückt eine andere Dimension in den Vordergrund: die reflexive Erfahrung. Der hier vorgeschlagene Begriff einer *Ethik der reflexiven Erfahrung* bezeichnet jene Haltung, die nicht aus theoretischem Vorwissen oder methodischer Souveränität hervorgeht, sondern aus dem gelebten Durchgang durch eigene psychische Prozesse in ihrer existenziellen Bedeutung. Erfahrung ist hier also nicht im Sinne empirischer Datenerhebung verstanden, sondern als reflexiv erschlossene Erfahrung des eigenen psychischen Geschehens im therapeutischen Vollzug. Eine solche Ethik zeigt sich u. a. in der Fähigkeit, Ambivalenz auszuhalten, affektiv präsent zu bleiben und auch unter Bedingungen von Unsicherheit relationale Urteilskraft zu bewahren. Psychotherapeutisches Wissen entsteht nicht ausserhalb des therapeutischen Geschehens, sondern im Prozess der therapeutischen Begegnung. Es ist gebunden an Zeit, Beziehung und Affekt sowie an die Bereitschaft, sich von dem berühren zu lassen, was sich im therapeutischen Raum zeigt. In diesem Sinn ist Wissen in der Psychotherapie keine primäre, sondern eine *abgeleitete* Kategorie. Es erhält seine Bedeutung erst dort, wo es sich im Kontakt mit subjektiver Erfahrung bewährt, korrigieren lässt oder an seine Grenzen stösst.

Diese Perspektive findet Anschluss an psychoanalytische Positionen, die das Aushalten von Nicht-Wissen als zentrale professionelle Fähigkeit begreifen. Wilfred Bion (1992) betont mit seinem Konzept der *negative capability*, dass vorschnelles Verstehen häufig weniger Ausdruck von Erkenntnis als von Abwehr ist. Verstehen erscheint hier nicht als Aneignung, sondern als Prozess, der Zeit benötigt und Unsicherheit einschliesst. Auch Donald Winnicott (2020) hebt hervor, dass therapeutische Wirksamkeit nicht aus Technik erwächst, sondern aus der Fähigkeit, einen haltenden Raum bereitzustellen, in dem Unreifes, Widersprüchliches und noch Nicht-Formuliertes existieren darf. Einen frühen psychoanalytischen Impuls in diese Richtung formulierte bereits Sándor Ferenczi (1932), der auf die Wirksamkeit nicht reflektierter Anteile aufseiten der Therapeut*innen hingewiesen hat. Seine Kritik an der Illusion analytischer Neutralität macht deutlich, dass Wissen dort problematisch wird, wo es zur Abwehr affektiver Involviertheit eingesetzt wird. Erfahrung fungiert hier nicht als Ergänzung zum Wissen, sondern als dessen notwendige Korrektur. Aktuelle relationale und intersubjektive Ansätze führen diese Perspektive fort, indem sie psychotherapeutisches Verstehen zunehmend als reflexiven und wechselseitig involvierten Prozess begreifen (vgl. Orange, 2011).

Aus psychotherapiewissenschaftlicher Perspektive wird diese Spannung zwischen Wissen und Nicht-Wissen auch von Hans-Joachim Hannich (2023, S. 62ff.) explizit thematisiert. Er beschreibt psychotherapeutisches Verstehen als ein Spannungsfeld, das per se konfliktthaft ist und dessen grundlegende Widersprüche nicht auflösbar sind. Die damit verbundene Unsicherheit und Ungewissheit müssen toleriert werden, ebenso die Affekte, die aus dieser Widersprüchlichkeit hervorgehen. Erkenntnis setzt hier nicht Auflösung voraus, sondern die Fähigkeit, das Spannungsfeld von Wissen und Nicht-Wissen auszuhalten. Gerard Burda (2025) radikalisiert diese erkenntnistheoretische Problematik, indem er sowohl realistische als auch konstruktivistische Positionen als Ausdruck zugrunde liegender, nicht weiter begründbarer Grundannahmen beschreibt und daraus eine prinzipielle Begrenztheit wissenschaftlicher Erkenntnis ableitet. Sein Konzept des unverfügbaren Subjekts verweist auf eine Dimension, die sich weder vollständig in bestehende Ordnungssysteme integrieren noch aus ihnen ausschliessen lässt. Der hier vertretene Ansatz teilt diese erkenntniskritische Einsicht, gelangt jedoch zu einer anderen Schlussfolgerung hinsichtlich ihrer wissenschaftspraktischen Konsequenzen. Die Begrenztheit epistemischer Zugriffe führt nicht zur Infragestellung psychotherapiewissenschaftlicher Erkenntnis, sondern bestimmt deren spezifische Form. Psychotherapiewissenschaft operiert nicht im Modus endgültiger Objektivierung, sondern im Modus reflexiven Verstehens.

In diesem Zusammenhang erscheint die von Greiner (2024b, S. 33) entwickelte zweite Funktionsebene psychotherapiewissenschaftlicher Reflexion (FE2) als entscheidender methodischer Schritt. Sie bezeichnet jene Ebene, auf der psychotherapiewissenschaftliche Erkenntnis nicht nur ihren Gegenstand untersucht, sondern ihre eigenen begrifflichen Voraussetzungen, impliziten Vorannahmen und erkenntnistheoretischen Bedingungen selbst zum Gegenstand der Reflexion macht. In einem weiteren Sinne lässt sich diese reflexive Bewegung auch mit Romanyszyn (2021, S. 222) in Verbindung bringen, der wissenschaftliches Verstehen als einen Prozess beschreibt, in dem die oft unbewussten Dynamiken des eigenen wissenschaftlichen Arbeitens zunehmend bewusst werden. Erkenntnis entsteht in dieser Perspektive nicht ausserhalb, sondern innerhalb eines reflexiven Erkenntnisprozesses. Psychotherapiewissenschaft erscheint damit nicht trotz, sondern aufgrund der Unabschliessbarkeit ihres Gegenstands als eigenständige wissenschaftliche Praxis.

Erfahrungsräume einer Ethik der reflexiven Erfahrung

Wenn die *Ethik der reflexiven Erfahrung* nicht als abstraktes Prinzip missverstanden werden soll, sondern als gelebte Haltung psychotherapeutischer Praxis, kommt Lehrtherapie und Supervision eine zentrale Bedeutung zu. Sie sind nicht bloss Ausbildungsbestandteile oder Instrumente der Qualitätssicherung, sondern relationale Erfahrungsräume, in denen sich jene epistemische Ethik ausbildet, die psycho-

therapeutisches Handeln vor Reduktion, *Vereindeutigung* und Machtmissbrauch schützt. In ihnen wird nicht primär vermittelt, *was* zu tun ist, sondern erfahrbar, *wie* Wissen, Beziehung und Begrenztheit ineinandergreifen.

Lehrtherapie konfrontiert angehende Psychotherapeut*innen mit der eigenen Subjektivität, mit Abhängigkeit, Ungeduld und der Erfahrung, nicht vollständig souverän zu sein. Dadurch unterläuft sie die Vorstellung, psychotherapeutische Kompetenz könne primär aus Wissen und Technik abgeleitet werden. Verstehen erweist sich hier nicht als verfügbarer Besitz, sondern als prozesshaftes Geschehen, das sich erst im Durchgang durch Beziehung ausbildet. *Supervision* erweitert diesen Erfahrungsraum um eine kontinuierliche reflexive Dimension. Sie macht das eigene therapeutische Handeln selbst zum Gegenstand der Beziehung und exponiert blinde Flecken, implizite Annahmen und affektive Verstrickungen. Auch hier geht es nicht um die Herstellung richtiger Lösungen, sondern um die Fähigkeit, das eigene Tun im Spannungsfeld von Verantwortung, Unsicherheit und Beziehung zu reflektieren. Supervision wirkt damit als dauerhaftes Korrektiv gegen epistemische Hybris und gegen die Versuchung, Wissen zur Absicherung von Macht zu verwenden. Beide Formate verankern psychotherapeutisches Wissen im Erleben. Sie machen erfahrbar, dass professionelle Haltung nicht aus Überlegenheit entsteht, sondern aus der Bereitschaft, sich korrigieren, irritieren und begrenzen zu lassen. Lehrtherapie und Supervision sind damit keine nachträglichen Ergänzungen eines ansonsten vollständigen Kompetenzmodells, sondern dessen Voraussetzung.

Problematisch werden sie dort, wo sie im Zuge zunehmender Akademisierung marginalisiert, verkürzt oder funktionalisiert werden. Sobald Lehrtherapie und Supervision auf überprüfbare Lernziele, standardisierte Abläufe oder formale Nachweise reduziert werden, verlieren sie ihre ethische Wirksamkeit. Die *Ethik der reflexiven Erfahrung* widersetzt sich dieser Instrumentalisierung, indem sie anerkennt, dass zentrale Dimensionen professioneller Kompetenz – wie Ambiguitätstoleranz, affektive Präsenz und relationale Urteilskraft – sich einer vollständigen Operationalisierung entziehen. Lehrtherapie und Supervision halten damit jene Spannung aufrecht, die für psychotherapeutische Praxis konstitutiv ist: zwischen Wissen und Erfahrung, Nähe und Distanz, Verantwortung und Begrenztheit. Wo diese Spannung nicht mehr ausgehalten wird, verliert Psychotherapie ihren ethischen Boden. In diesem Sinn sind Lehrtherapie und Supervision nicht nur Ausbildungsformate, sondern Schutzräume des Psychotherapeutischen selbst.

Psychotherapie und das Unauflösbare

Psychotherapeutische Praxis bewegt sich im Spannungsfeld zwischen dem, was verstanden, gelindert oder verändert werden kann, und dem, was sich einer solchen Bearbeitung entzieht. Leiden ist dabei nicht in jedem Fall Ausdruck von Pathologie, sondern häufig Teil menschlicher Existenz – verwoben mit Endlichkeit, Verlust,

Ambivalenz und Beziehung. Diese Grundannahme hat nicht nur praktische, sondern auch erkenntnistheoretische Konsequenzen. In manchen Situationen lässt sich Leiden einordnen oder transformieren. In anderen bleibt es als Erfahrung bestehen, die sich nicht auflösen lässt, sondern nur im Vollzug von Beziehung tragbar wird. Psychotherapie operiert damit nicht ausschliesslich im Modus gezielter Veränderung, sondern ebenso in einem Modus des Begleitens, in dem Verstehen nicht als Auflösung, sondern als Annäherung verstanden wird. Erkenntnis vollzieht sich hier prozessual, situativ und relational – und entzieht sich damit einer vollständigen Fixierung.

Ein epistemisches Risiko entsteht dort, wo dieses Spannungsfeld zugunsten von Eindeutigkeit und Lösbarkeit verkürzt wird. Wird Leiden primär als Defizit konzeptualisiert, das erkannt und behoben werden muss, verschiebt sich der Fokus vom verstehenden Umgang mit subjektiver Erfahrung hin zu ihrer diagnostischen Erfassbarkeit und funktionalen Bearbeitbarkeit. Psychotherapie gerät dann in Gefahr, ihre eigene erkenntnistheoretische Voraussetzung zu verfehlen, indem sie Komplexität reduziert, wo Offenheit erforderlich wäre. In dieser Perspektive wird deutlich, dass epistemische *Macht* nicht erst dort wirksam wird, wo explizit normiert oder kontrolliert wird, sondern bereits dort, wo festgelegt wird, was als relevantes Erleben, angemessene Entwicklung oder gültige Bedeutung gilt. Der Anspruch, das Erleben des Anderen eindeutig erfassen zu können, unterläuft die Anerkennung seiner Subjektivität als eigenständige Erkenntnisquelle. Psychotherapeutisches Wissen verliert an diesem Punkt seine dialogische Struktur.

Die *Ethik der reflexiven Erfahrung* markiert hier zugleich eine epistemische *Grenze*. Sie verweist auf Dimensionen menschlichen Erlebens, die sich einer vollständigen begrifflichen Erfassung entziehen und nur im zeitlichen Vollzug von Beziehung zugänglich werden. Leiden, Ambivalenz und biografische Brüchigkeit erscheinen nicht als Defizite des Wissens, sondern als Hinweise auf dessen Begrenztheit. Erkenntnis zeigt sich hier nicht im Besitz von Wahrheit, sondern im Aushalten von Unabschliessbarkeit. Psychotherapeutisches Handeln bedeutet aus dieser Sicht, einen Erkenntnisraum zu eröffnen, in dem Widerspruch, Mehrdeutigkeit und Vorläufigkeit bestehen dürfen. Verstehen wird nicht als abschliessendes Deuten verstanden, sondern als gemeinsamer Prozess, der offen bleibt für Revision, Irritation und Veränderung. Das Unauflösbare ist dabei kein Scheitern von Erkenntnis, sondern ihr konstitutiver Horizont. Aus dieser Haltung ergibt sich die epistemische Bedeutung von Pluralität. Wo menschliches Leiden nicht vollständig erklärbar oder in einer einzigen Perspektive erschöpfend verstehbar ist, kann es keine singuläre, abschliessende Form des Verstehens geben. Pluralität erscheint daher nicht als Mangel an Systematik, sondern als Ausdruck epistemischer Bescheidenheit und wissenschaftlicher Reife. In diesem Sinn lässt sich mit Safranski (2009, S. 78) sagen, dass die Einheit der Vernunft nicht in Vereinheitlichung besteht, sondern in der Vielfalt ihrer Stimmen. Für die Psychotherapiewissenschaft bedeutet dies, Differenz nicht aufzulösen, sondern als Bedingung von Erkenntnis reflexiv auszuhalten.

Zur Notwendigkeit einer eigenständigen Psychotherapiewissenschaft

Diese Überlegungen machen deutlich, dass die ethische Dimension psychotherapeutischer Praxis untrennbar mit ihrer wissenschaftlichen Verortung verbunden ist. Wird Psychotherapie als Praxis des Begleitens, Verstehens und Aushaltens des Unauflösbaren ernst genommen, kann sie nicht in Wissenschaftslogiken aufgehen, die primär auf Erklärung, Standardisierung und Kontrolle ausgerichtet sind. Die hier entwickelte Ethik der reflexiven Erfahrung verweist vielmehr auf die Notwendigkeit eines wissenschaftlichen Rahmens, der ihrer erkenntnistheoretischen Eigenart entspricht. In diesem Sinn ist die Forderung nach einer eigenständigen Psychotherapiewissenschaft keine disziplinpolitische Abgrenzung, sondern eine Konsequenz aus dem Gegenstand der Psychotherapie selbst. In der Tradition Greiners lässt sich Psychotherapie als eigenständige Wissenschaft begreifen, die sich hermeneutisch, relational und modalitätenübergreifend konstituiert und deren Erkenntnisinteresse auf den verstehenden Umgang mit Sinn- und Bedeutungszusammenhängen gerichtet ist. Ihre Aufgabe besteht nicht darin, Modelle benachbarter Disziplinen zu übernehmen, sondern eine eigene epistemische Logik zu entwickeln, die subjektivem Erleben, Prozessualität und Intersubjektivität gerecht wird.

Diese Ethik kann in diesem Zusammenhang als ethische Zuspitzung psychotherapiewissenschaftlichen Denkens verstanden werden. Sie begreift Psychotherapie nicht als Technik der Lösung, sondern als relationale und verantwortliche Praxis und markiert jene Grenze, an der Wissen seine souveräne Position verliert und sich als abhängig von Erfahrung, Beziehung und Zeit erweist. Psychotherapiewissenschaft übernimmt damit nicht nur eine erkenntnistheoretische, sondern auch eine ethische Verantwortung: das Subjektive vor Reduktion zu schützen und Psychotherapie nicht zur Praxis der *Vereindeutigung* werden zu lassen. Gerade im Kontext zunehmender Akademisierung wird deutlich, dass Psychotherapie nicht bruchlos in bestehende Wissenschaftslogiken integriert werden kann, ohne ihre eigenen Voraussetzungen zu verfehlen. Eine eigenständige Psychotherapiewissenschaft wirkt dieser Gefahr entgegen, indem sie Pluralität, Ambiguität und das Unauflösbare nicht als Defizite, sondern als konstitutive Dimensionen menschlicher Erfahrung anerkennt und bewahrt.

Literatur

- Bauer, T. (2018). *Die Vereindeutigung der Welt. Über den Verlust an Mehrdeutigkeit und Vielfalt*. Reclam.
- Bion, W.R. (1992). *Lernen durch Erfahrung*. Übers. v.E.Krejci. Suhrkamp.
- Burda, G. (2025). *Das unverfügbare Subjekt: Ein psychotherapiewissenschaftsphilosophisches Manifest für das, was weder ein- noch ausgeschlossen werden kann*. Waxmann.
- Fiedler, P. (2012). *Die Zukunft der Psychotherapie: Wann ist endlich Schluss mit der Konkurrenz?* Springer.
- Fischer, G. et al. (2009). *Warum eine eigenständige Psychotherapiewissenschaft dringend gebraucht wird: Gegen Trivialisierung und Bildungsverlust der Psychotherapie*. Asanger.

- Ferenczi, S. (1932). Confusion of Tongues Between Adults and the Child. In *Final Contributions to the Problems and Methods of Psycho-Analysis*. Hogarth.
- Grawe, K. et al. (1994). *Psychotherapie im Wandel: Von der Konfession zur Profession*. Hogrefe.
- Greiner, K. (2015). Akademische Psychotherapie. Philosophische Fundierung einer neuen Einzelwissenschaft an der Sigmund-Freud-Privatuniversität Wien (SFU). In K. Greiner & M. Jandl (Hg.), *Bizarrosophie. Radikalkreatives Forschen im Dienste der akademischen Psychotherapie* (S. 11–34). Traugott Bautz.
- Greiner, K. (2017). Wie man Poppers philosophischen Knüppel in einen Blumenstrauß für die Psychoanalyse verwandelt: Ein psychotherapiewissenschaftstheoretischer Essay. *Psychotherapie-Wissenschaft*, 7(2), 77–83.
- Greiner, K. (2022). Wieso Wirkamkeitsstudien nichts über den Wissenschaftlichkeitsstatus von Psychotherapiemodalitäten aussagen können. *SFU Forschungsbulletin*, 10(1), 86–90.
- Greiner, K. (2024a). Psychotherapie als Textmedizin. Versuch über ein allgemeines Funktionsparadigma. *Psychotherapie-Wissenschaft*, 14(1), 9–17.
- Greiner, K. (2024b). *Lehrbuch Experimentelle Psychotherapiewissenschaft. Innovative Therapieschulenforschung an der SFU Wien seit 2007*. SFU Wien.
- Greiner, K. (2025). Akademische Psychotherapie. Epistemologische Begründung eines neuen Universitätsfachs. In I.M. Reményi & C. Ahlers (Hg.), *Mehtaverse. Festschrift für Gerda Mehta* (S. 303–323). SFU Wien.
- Hannich, H.J. (2023). Der wissenschaftliche Zugang der Hermeneutik zur Psychotherapie. In S. Sulz (Hg.), *Von der Psychotherapiewissenschaft zur Kunst der Psychotherapie*. Psychosozial-Verlag.
- Jandl, M. & Greiner, K. (2023). Logopoietische Hermeneutik als Gegenwissenschaft im Konzert der humanwissenschaftlichen Psychotherapieforschung. In R. Popp et al. (Hg.), *Heilen – Bilden – Forschen – Managen. Im Mittelpunkt der Mensch. Festschrift für Alfred Pritz* (S. 141–152). Pabst.
- Kuhn, T.S. (1976). *Die Struktur wissenschaftlicher Revolutionen*. 2., rev. u.erg. Aufl. Suhrkamp.
- Orange, D.M. (2011). *The suffering stranger: Hermeneutics for everyday clinical practice*. Routledge.
- Pritz, A. & Teufelhart, H. (1996). Psychotherapie – Wissenschaft vom Subjektiven. In A. Pritz (Hg.), *Psychotherapie – eine neue Wissenschaft vom Menschen* (S. 1–18). Springer.
- Raile, P. (2023). *Psychotherapiewissenschaft. Grundlagen einer eigenständigen wissenschaftlichen Disziplin*. Waxmann.
- Romanyshyn, R.D. (2021). *The wounded researcher: Research with soul in mind*. Routledge.
- Safranski, R. (2009). Selbstverhältnisse, Weltverhältnisse. In M. Jongen et al. (Hg.), *Die Vermessung des Ungeheuren. Philosophie nach Peter Sloterdijk* (S. 73–79). Fink.
- Sluneko, T. & Benetka, G. (2023). *Geschichte und Paradigmen der Psychologie und Psychotherapie*. Facultas.
- Smith, J. (2025). Psychotherapy integration from the bottom up: A unifying, science-based view of psychotherapy's infrastructure. *Journal of Psychotherapy Integration*, 35(4), 256–271.
- Winnicott, D.W. (2020). *Reifungsprozesse und fördernde Umwelt: Studien zur Theorie der emotionalen Entwicklung*. Übers. v.G. Theusner-Stampa. Psychosozial-Verlag.

On the epistemic autonomy of the science of psychotherapy

A plea for an ethics of reflexive experience

Abstract: The increasing academization of psychotherapy raises fundamental questions about its scientific positioning and ethical foundations. In particular, trends toward standardization, homogenization, and psychologization threaten to marginalize those

dimensions that are constitutive of psychotherapeutic practice: subjective experience, processuality, ambiguity, and the irreducible nature of human suffering. his article argues that psychotherapy follows its own epistemic logic, one that cannot be subsumed under explanatory or technicist models of science. Against this backdrop, the article makes a case for an ethics of reflexive experience as a central prerequisite for the epistemic autonomy of the science of psychotherapy. Knowledge is not understood here as a primary entity, but rather as a derived category that is legitimized only through its engagement with reflexively explored experience, relationships, and time. Training therapy and supervision are identified as central spaces of experience in which this ethical stance is developed and safeguarded against epistemic hubris. This text is intended as a contribution to the self-understanding of the science of psychotherapy and as a counterpoint to standardizing tendencies in current scientific discourse.

Keywords: psychotherapy science, ethics of reflexive experience, epistemic autonomy, subjective experience, intersubjectivity, not-knowing, tolerance of ambiguity, the irreducibility of human suffering, psychotherapy

Biografische Notiz

Tamara M. Trebes ist Psychoanalytikerin und Psychotherapiewissenschaftlerin in Wien. Sie arbeitet in freier Praxis, ist Dissertantin an der Sigmund Freud PrivatUniversität (SFU) und lehrt sowohl an der SFU als auch an der Musik und Kunst Privatuniversität der Stadt Wien (MUK).

Kontakt

E-Mail: Tamara.Trebes@web.de

Buchbesprechungen

Geuter, U. (2026). *Wege zum Körper. Schriften zur Körperpsychotherapie*. Psychosozial-Verlag, 364 S., 61.90 CHF, 44.90 EUR, ISBN: 978-3-8379-3484-7

Psychotherapie-Wissenschaft 16 (2) 2026 63–64

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-63>

Ulfried Geuter ist aktuell einer der renommiertesten Vertreter der Körperpsychotherapie im deutschen Sprachraum. Viele Leser werden ihn als Autor kennen der beiden Bände *Körperpsychotherapie. Grundriss einer Theorie für die klinische Praxis* (2023, 2. Aufl.) und *Praxis Körperpsychotherapie* (2019), beide bei Springer erschienen. Sie gelten bereits heute als Standardwerke der Körperpsychotherapie.

Das vorliegende Buch in Form eines Sammelbands umfasst insgesamt 15 Kapitel. 13 hatte der Autor zwischen 1996 und 2023 bereits an anderer Stelle veröffentlicht. Die einzelnen Kapitel stehen für sich selbst. Man kann als Leser seinen eigenen Interessen folgen. Das Buch wird den Neuling auf dem Gebiet der Körperpsychotherapie motivieren, sich weiter mit dem Körper in der Psychotherapie zu beschäftigen. Der erfahrene Körperpsychotherapeut wird Neues erfahren. Geuter verfügt über ein enzyklopädisches Wissen. Seine ausführlich zitierte Literatur ist bereits für sich genommen ein wertvoller Schatz und verweist bei den verschiedenen Themen in die Tiefe. Er ist ein begabter Schreiber, seine Sprache ist eingängig und seine Texte lesen sich leicht. Behandeln die Kapitel Themen in abgeschlossener Form, erzeugt deren Reihenfolge eine gewisse Dramaturgie. Der Titel des Bandes *Wege zum Körper* ist mehrdeutig. Er nimmt den Leser gewissermaßen auf eine Reise mit. Es wird die Entstehungsgeschichte der Körperpsychotherapie in mehreren Kapiteln dargestellt, in sich ändernder Perspektive einer Spirale ähnelnd, die sich weiter hochdreht. Bis zu Kapitel 7 beschreibt Geuter mehrfach, dass die Körperpsychotherapie drei Traditionen in sich vereinigt. Die neoreichianische, die den Körper als Charakterpanzer sieht, und die darin fixierte Abwehr; die zweite, die auf Elsa Gindler und Reformgymnastik Anfang des 20. Jahrhunderts zurückgeht, bei der es um das körperliche Wahrnehmen, das Spüren und das Sich-nach-innen-Wenden geht; schliesslich die dritte, die beziehungsorientierte, die sich auf Ferenczi beruft.

Seine historische Versiertheit zeigt sich in Kapitel 4 über die Hintergründe zur Person des Psychotherapeuten Graf Dürkheim, dem Begründer der Initiatischen Leibtherapie. Geuter belegt, dass dieser mit dem Naziregime eng kollaboriert und sich nach dem Krieg nicht davon distanziert hat.

In Kapitel 8 legt Geuter dar, dass sich nach seinem Verständnis die Körperpsychotherapie nicht als Weiterführung einer tiefenpsychologischen Methode begreift. Sie hat sich vielmehr davon emanzipiert. Sie ist als eigenständige Richtung mit eigenem Profil zu verstehen, die der humanistischen Psychotherapie zuzurechnen ist. Diese Position unterlegt er in den Kapiteln 6–8. Für ihn haben die Säuglings- und Bindungsforschung sowie neurowissenschaftliche Resultate, z. B. die Entdeckung der Spiegelneuronen, und die Resultate der Emotionsforschung den körperbezogenen Ansatz objektivieren können. Diese Ergebnisse haben das bereits klinisch vorbestehende Wissen und die Praxeologie der Körperpsychotherapie legitimiert. Insbesondere ist für Geuter George Downings Theorie der affektmotorischen Schemata wertvoll. Das Verständnis von Embodiment, wie Geuter in Kapitel 7 darlegt, stellt das Fundament dar für seinen erfahrungszentrierten Ansatz. Die im Hier und Jetzt entstehende Therapiesituation zwischen Patient und Therapeut ist zum einen eine Co-Kreation zwischen beiden und beruht zum anderen auf einer zwischen Körper und Geist untrennbaren, holistischen Erfahrung. In Kapitel 8 stellt er das Selbsterleben ins Zentrum der Psychotherapie und fasst dies als ein zutiefst körperliches Erleben auf, bestehend aus extero-, proprio- und interozeptiver Wahrnehmung. Der Leser erhält dadurch Einsicht in seinen eigenen Prozess, den der Autor im Verlauf der Jahre gemacht hat. Mehrfach kommentiert Geuter das damals Geschriebene in Fussnoten selbstkritisch aus heutiger Sicht. Das macht ihn sympathisch und gleichzeitig zeugt es von seinem wachen Geist.

In weiteren Kapiteln wendet er sich klinisch-praktischen Fragen zu und unterlegt sie jeweils mit kurzen Fallgeschichten. In Kapitel 11 führt er sein Verständnis von der Arbeit mit Emotionen aus, wie bei blockierter oder fehlender emotionaler Anspannung, bzw. bei zu heftiger emotionaler Anspannung, vorzugehen ist. In Kapitel 12 beleuchtet er den Stellenwert eines achtsamkeitsorientierten Ansatzes, bzw. deren Art der Interventionen, und wie dieser in eine erlebniszentrierte Körperpsychotherapie einfließen kann. In Kapitel 13 gibt er einen hervorragenden Kurzüberblick über die verschiedenen Formen von Traumata und Traumafolgestörungen, wie sich diese auf das Körpererleben auswirken und welche Möglichkeiten des körperpsychotherapeutischen Zugangs bestehen. Er integriert dabei verschiedene Ansätze, wie Somatic Experiencing nach Levine, die Screen-Technik und EMDR. Geuter hebt die Wichtigkeit der Beachtung der drei Phasen hervor von Stabilisierung, schonender Konfrontation, die kontrolliert und begrenzt erfolgen muss, und der dritten Phase, in der das Erlebte reflektiert, verstanden und eingeordnet werden muss. Im klinischen Beispiel zeigt er, wie er das EMDR-Standardprotokoll auf einen körperorientierten Ansatz abwandelt.

Kritisch äussert sich Geuter zu Porges Vagaltheorie. Einerseits würdigt er dessen klinische Anschaulichkeit mit den drei unterscheidbaren Reaktionen auf Bedrohung: Aktivierung des Bindungssystems, der Flucht oder des Kampfs und bei Überwältigung mit Immobilisierung. Er hebt aber hervor, dass die bei vielen Psychotherapeuten verbreitete und beliebte Überzeugung einer neurowissenschaftlichen Evidenz, von Porges vertreten, nicht haltbar ist. Weder neuroanatomische noch neurophysiologische Untersuchungen bestätigen Porges Behauptungen, dass eine eindeutige Zuordnung psychischer Zustände zu bestimmter Nerventätigkeit belegbar ist.

Im 15. Kapitel fasst er seinen erlebniszentrierten körperpsychotherapeutischen Ansatz zusammen, der keiner der aktuell bestehenden körperpsychotherapeutischen Schulen zuzuordnen ist. Diesen hat er in seinem 2019 publizierten Buch detailliert dargestellt. Es sind allgemeingültige, insgesamt 10 Prinzipien der klinischen Praxis. Sie legen dar, was mit körperorientierten Mitteln bewirkt werden kann.

Das Buch ist für eine breite Leserschaft mit unterschiedlichen Vorkenntnissen bereichernd und lesenswert, auch wenn man nicht immer der gleichen Meinung sein muss. In Kapitel 9 geht Geuter der Frage nach, ob das Konzept der Lebensenergie für die Körperpsychotherapie klinisch sinnvoll ist, und begründet ausführlich seine Position, dass darauf verzichtet werden kann. Was vermissen wir in diesem Band? In der Körperpsychotherapie ist die Frage nach dem Wesen der Aggression zentral. Bei Lowen etwa nimmt sie einen zentralen Stellenwert ein: Er vertritt die Auffassung, dass sie etwas Gesundes ist, der besseren Abgrenzung und Selbstbehauptung dient und dass die Erfahrung des Grounding hier hilfreich ist. Im gesamten Band wird Aggression dagegen nur an einer Stelle erwähnt – Reich bewertet sie, ebenso wie Lowen, als positiv. Es fehlt eine fundierte Auseinandersetzung mit dem Thema, auch mit der Frage, ob es eine primäre Aggression gibt, die bei Fehlleitung destruktiv werden kann. In diesem Fall wird der Körper zum Werkzeug der Destruktion.

Geuters Horizont ist breit und tief; entsprechend gründlich recherchiert er. Als inhaltlicher Fehler fällt einzig auf, dass er behauptet, dass die Gesichtsmuskulatur vom Nervus vagus als Teil des Parasympathikus innerviert bzw. aktiviert wird. Das ist ein von Porges behauptetes Postulat. Anatomisch ist das unzutreffend. Es ist der Nervus facialis mit seinen speziell-viszeromotorischen Fasern, die für die Innervation der gesamten Gesichtsmuskulatur verantwortlich sind. Sie sind nicht Teil des Parasympathikus. Hier sitzt Geuter Porges auf. Porges postuliert eine funktionelle/regulatorische Kopplungshypothese verschiedener Hirnnerven. Der Nervus facialis bewegt jedoch selbst die Muskeln – der Vagus «spricht» nur mit benachbarten Hirnstammzentren.

Die Lektüre bereichert und regt zur Vertiefung einzelner Themen an. Das Buch ist uneingeschränkt empfehlenswert.

Thomas Fellmann

Stöckli, P. (2026). *Autismus, Trauma und Bindung. Neue Wege zu Regulation und Verbindung*. Kohlhammer, 269 S., 55.90 CHF/46.80 CHF, 39.– EUR, ISBN: 978-3-17-046437-7

Psychotherapie-Wissenschaft 16 (2) 2026 64–65

www.psychotherapie-wissenschaft.info

CC BY-NC-ND

<https://doi.org/10.30820/1664-9583-2026-2-64>

Philippe Stöckli legt ein Buch vor, das die Therapie von Menschen im Autismus-Spektrum verändern und befruchten wird. Er reflektiert die wechselseitigen Wirkungen von Autismus, Trauma und Bindung. Das Buch hat einen gut strukturierten Aufbau. Der Autor zeigt zu Beginn auf, dass es bisher kein klares Autismus-Konzept gibt, und fordert ein solches ein. Dieses Buch soll einen Beitrag dazu leisten. So referiert er das bisherige Autismus-Verständnis und die vorliegenden Theorien. Autismus wurde als unveränderliche genetische Disposition verstanden und die Therapie zielte auf Akzeptanz dieses unabänderlichen Andersseins und auf Training von sozialen Kompetenzen oder Energiemanagement. Dann erläutert der Autor den Zusammenhang von Autismus und Trauma, Trauma und Stress sowie Autismus und frühkindliche Bindungsentwicklung, worauf er ein Vier-Komponenten-Modell entwickelt, das die Basis für ein integratives Behandlungsmodell liefert. Der Einbezug von Behandlungsansätzen aus der Traumatherapie bietet neue Entwicklungschancen auch für Autismus-PatientInnen.

Stöckli integriert Erkenntnisse aus der Psychotraumatologie, der Bindungstheorie, der Stressforschung und der Neurobiologie. Die Fülle rezipierter Literatur aus all diesen Gebieten ist beeindruckend. Auf diese Weise ersetzt er die bisherigen Symptomkategorien durch eine funktionale Kategorisierung, die auf vier zugrundeliegenden Pathomechanismen beruht. Der therapeutische Ansatz berücksichtigt Aspekte wie Informationsintegration, Regulation, Stress und Trauma, Bindung und Körper (Verbindung nach innen und aussen) sowie Kompensations- und Kontrollstrategien.

Die genetische Komponente kann mit Psychotherapie nicht therapiert werden, sie gilt es anzunehmen. Im Bereich der Stress- und Trauma-, der Bindungs- und der Kompensationskomponente eröffnen sich jedoch viele Möglichkeiten, psychotherapeutisch Veränderungen und Wachstum zu fördern. Viele Menschen im ASS-Spektrum haben Traumatisierungen erlebt. Diese gilt es zu behandeln. So berücksichtigt Stöcklis Behandlungsmodell die Herstellung von Sicherheit und Ernst-genommen-Sein als Voraussetzung einer Therapie sowie eine Förderung der Innenorientierung durch Achtsamkeit auf eigene Körperwahrnehmungen. Dann gilt es, eine sichere Bindung zu fördern. Akzeptanzorientierte Elemente wollen mit einer therapeutischen Haltung, wie sie den humanistischen Therapieformen eigen ist (bedingungslose Wertschätzung, Authentizität, Empathie), die Scham des Andersseins reduzieren. Aus traumaorientierten Verfahren ist bekannt,

dass mit einer Verbesserung der Körperwahrnehmung auch die Wahrnehmung von Emotionen gefördert wird. Das gilt es auch bei ASS-PatientInnen zu fördern. So können Gefühle besser wahrgenommen und identifiziert werden. Ziel einer Traumatherapie ist eine Integration von überwältigenden Inhalten. Durch Lösung von Energieblockaden kann eine Person aus der Erstarrung «erwachen». Dabei gilt es, eine Haltung des respektvollen Grenzenwahrens und Explorierens einzunehmen. Beim körperorientierten Arbeiten empfiehlt der Autor Somatic Experiencing.

Abschliessend fasst Stöckli die zentralen Thesen des Buchs zusammen. Ein Grossteil der Autismus-Symptome würden direkte oder indirekte Reaktionen aus Überforderungssituationen darstellen: «Durch kleinschrittige Wiederherstellung einer gesunden Widerstandskraft, Abgrenzung, Filterung der Einflüsse und Energiemanagement können sich die inneren Bedingungen so verändern, dass Integration und Regulation wieder möglich sind und ein Gleichgewicht hergestellt wird» (S. 244). Wohltuend und in wissenschaftlichen Publikationen selten anzutreffen sind auch seine kurzen Ausführungen zu seinem persönlichen Hintergrund und seiner Motivation zu diesem Buch. Ganz im Sinne Paul Goodmans.

Gern empfehle ich dieses reichhaltige Buch zur Lektüre. Es richtet sich nicht nur an ein Fachpublikum, sondern etwa auch an Eltern, die besser verstehen wollen, wie Störungen im Autismus-Spektrum sich erklären und behandeln lassen.

Peter Schulthess

Hell, D. (2026). *Fragile Psyche. Von Umbrüchen und Aufbrüchen*. Psychosozial-Verlag, 202 S., 35.90 CHF, 27.94 EUR, ISBN: 978-3-8379-3386-4

Psychotherapie-Wissenschaft 16 (2) 2026 65
www.psychotherapie-wissenschaft.info
 CC BY-NC-ND
<https://doi.org/10.30820/1664-9583-2026-2-65>

Daniel Hell ist als Zürcher Psychiater und Psychotherapeut vielen ein Begriff. In seinem neuesten Buch fokussiert er auf die Fragilität des Menschen. Er führt aus, dass Fragilität eine Lebenskonstante von uns Menschen ist, von Geburt bis Tod. Umbrüche gehören zu jedem Leben. Fragil und verletzlich zu sein, ist nicht eine Krankheit, sondern etwas ganz Normales. In systemati-

scher Weise geht er der Bedeutung und Wertschätzung von Fragilität seit dem Altertum in Philosophie, Gesellschaft und Psychotherapie nach. Beeindruckend ist die Wissensbreite und Belesenheit des Autors, die in diesem Buch zum Ausdruck kommt. Der Autor formuliert folgende Kernthesen:

Fragilität ist ein inhärenter Teil des Menschseins. Sie ist kein «Betriebsunfall» oder ein Zeichen von Labilität, sondern eine grundlegende Eigenschaft der menschlichen Psyche. Der Mensch ist insbesondere deshalb fragil, weil er die Fähigkeit zur Selbstreflexion hat. Diese ermöglicht ihm, sich selbst infrage zu stellen und zu bewerten, was zwar Freiheit schafft, aber auch eine permanente Angriffsquelle ist für Selbstzweifel, Scham und Leiden.

Das Leben verläuft nicht linear, sondern ist von Brüchen, Krisen und Umbrüchen durchzogen. Diese beinhalten einerseits ein Risiko für psychische Bruchereignisse, Fragmentierungen oder Erkrankungen (wie etwa Depressionen), bieten aber andererseits auch eine Chance für Wachstum, Lernerfahrungen und schöpferische Tätigkeit. Mit der Annahme der eigenen Zerbrechlichkeit können Krisen ins Leben integriert werden.

Kritik am Optimierungswahn und Pathologisierung: Hell warnt davor, Gefühle wie Scham oder Traurigkeit vorschnell als krankhaft abzutun und durch ständige Selbstoptimierung zu verdrängen. Gerade in unserer schnelllebigen Zeit voller gesellschaftlicher Veränderungen und wegen des Leistungsdrucks, immer mehr zu erreichen, kann eine Kluft zwischen Selbstbild und Realität entstehen. Scham ist ein guter Indikator für das Erkennen einer solchen Kluft. Traurigkeit hingegen kann ein Resilienzfaktor sein, um Verluste, Enttäuschungen und Krisen zu verarbeiten.

Der Autor beleuchtet zudem die Rolle der Fragilität in verschiedenen Lebensphasen, von den ersten Bindungserfahrungen nach der Geburt über die Adoleszenz bis hin zum Alter.

Hell plädiert für die Annahme der eigenen Endlichkeit und Fragilität und für einen konstruktiven Umgang damit. «Risse im Selbstbild» seien nicht zu leugnen, sondern als notwendige Bedingung für Wachstum zu akzeptieren. Dem Autor gelingt es, in flüssiger Sprache wissenschaftliche und philosophische Erkenntnisse mit seiner klinischen Erfahrung zu verbinden, die er mit Fallbeispielen illustriert. Das Buch ist geprägt von Daniel Hells eigener Menschlichkeit und seinem humanistischen Menschenbild. Ich empfehle es gern zur Lektüre.

Peter Schulthess

Kongresskalender

22.–25.10.2026

EAGT Konferenz

European Association for Gestalt Therapy
Bridges of Difference
Dubrovnik, Kroatien
<https://eagt2026dubrovnik.com>

14.–18.04.2027

The Tao of Gestalt

1st International China Gestalt Conference
Shanghai, China
<https://gestalt2027shanghai.com>

7.–9.05.2027

D-A-C-H Gestalt Konferenz

Nichts tun und alles wird anders?
Die paradoxe Theorie der Veränderung
und ihre Zumutungen für moderne Ichs
Zürich, Schweiz
<https://netzwerk-gestalttherapie.ch>

20.–27.05.2027

SEPI International Conference

The Future of Psychotherapeutics:
Technology, Knowledge Acquisition,
and The «Self» of the Psychotherapist
Newport R.I., USA
<https://www.sepiweb.org>

24.–26.09.2027

Gestalt Therapy Research Conference

Working together in Gestalt Therapy Research –
Cooperating among Clinicians
and Research Groups
Rom, Italien
www.eagt.org